Linear Rational Expectations Models
For Dynamically Interrelated Variables

Lars Peter Hansen*
Carnegie-Mellon University

Thomas J. Sargent
University of Minnesota
Federal Reserve Bank of Minneapolis

Revised January 1980

Working Paper #: 135

PACS File #: 2940

The research described in this paper was supported by the Federal Reserve Bank of Minneapolis. The computer programming was performed by Ian Bain, Thomas Doan, and Paul O'Brien. Ian Bain made extensive and very helpful comments on an earlier draft.

*Lars Peter Hansen was a 1979 summer visitor at the Federal Reserve Bank of Minneapolis.

Introduction

This paper aims to develop procedures for the rapid numerical computation and convenient methematical representation of a class of multiple variable, linear stochastic rational expectations models. A variety of examples from this class of models can be imagined. These include versions of interrelated factor demand models formed /like Mortensen's [18] by blending the model of Nadiri and Rosen [20] with the adjustment cost models of Lucas [13, 14], Gould [4], and Treadway [26,27]; models of exhaustible resource extraction along the lines of Epple and Hansen [3]; and dynamic linear models of interrelated industries, such as the corn and hog industries. Our desires for rapid computation and convenient representation are motivated by practical considerations, since our ultimate goal is to devise methods for estimating multiple-variable, rational expectations models of time series by versions of the method of maximum likelihood. Rapid computation of equilibria for different points in the parameter space is required for inexpensive maximization of the likelihood function. Convenient mathematical representation is valuable for a closely related reason, since it is desirable to be able to differentiate the likelihood function analytically in as many directions as possible. For this reason, a goal of the paper is to get "as close as possible" to a closed form summarizing the (highly nonlinear) cross-equation restrictions that are the hallmark of rational expectations models. It will become clear later what we mean by the phrase "as close as possible", since we shall indicate why a closed form analytic expression for the equilibrium cannot, in general, be calculated for the multiple endogenous variable models of the class that we consider.

This paper is a sequel to our earlier paper, Hansen and Sargent [7]. We extend the methods of estimating and formulating models that we have described in the single endogenous variable case to the case of several interrelated endogenous variables. We exploit several results obtained in the previous paper. Other important precursors to the work we do here include the work of Nadiri and Rosen [20], the paper by Lucas and Prescott [17] on investment, and the books of Holt, et al [8] and Graves and Telser [6]. Graves and Telser consider a certainty version of a problem close to our "standard problem". Our paper proposes a simple method for factoring the spectral-density-like matrix encountered in Graves and Telser's problem, and extends the computation of equilibria to the stochastic case by using a formula developed by Hansen and Sargent [7].

This paper is devoted solely to issues of model formulation. We have avoided issues of econometric estimation, including models or interpretations of "error terms", since these issues are extensively discussed in our earlier paper [7]. The estimation procedures and models of error terms described in that paper extend rather directly to the present class of setups. On the other hand, moving from a single endogenous variable to multiple variables does involve some nontrivial technical complications. It is these that we concentrate on in this paper.

The paper is organized as follows. Section 1 gives three examples of models that fall within the general class of models we study. The distinguishing characteristic of these models is that all are the solutions of quadratic dynamic optimum problems subject to linear constraints, and that to solve each one a spectral-density-like matrix must be factored. Section 2 describes the standard dynamic programming algorithm for solving our general problem, while sections 3 and 4 describe algorithms that are "faster" and "more revealing." Section 5 gives an example designed to illustrate the relative computation costs associated with the different algorithms. Section 6 then mentions a certain kind of "identification" or interpretation problem that can characterize some models of this class.

1. Three Examples

This section contains three examples of models that are included in the class of models analyzed in subsequent sections. The first two examples are related and are versions of interrelated factor demand models. We begin with the simpler of these models first, which is a model at the firm level and takes as given the factor price random processes and the output price random process. The second model then takes into account that, in the aggregate, firms' decisions influence output price. A third example indicates how the state and decision variables can be defined to model decisions about depletion of exhaustible resources.

Example (i): A firm's interrelated demand for factors

Letting E_0 be the mathematical expectation operator conditioned on information known at time 0, a firm is supposed to maximize

$$(1.1) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \left\{ P_t y_t - W_t n_t - J_t (k_t - k_{t-1}) - \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix}' D \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix} \right\}$$

subject to (k_{-1}, n_{-1}) given, and the linear production function

$$(1.2) \qquad y_t = d' \begin{bmatrix} k_t \\ n_t \end{bmatrix}$$

where d is a $2x1$ vector of positive constants. Here β is a discount factor with $0 < \beta < 1$, P_t is product price, k_t is the stock of capital, n_t the stock of employment of labor, W_t the wage rate of labor, J_t the price of capital goods, and D a positive definite matrix reflecting costs of adjusting factors of production. The model is a linear decision rule, quadratic objective function, multiple factor version of the costly adjustment models of Lucas [13,14], Treadway [26,27], Gould [4] and Mortensen [18]. The positive definite quadratic form in D represents interrelated costs of adjustment.

The firm maximizes (1.1) by choosing a contingency plan for setting $\begin{bmatrix} k_t \\ n_t \end{bmatrix}$, $t \geq 0$, as a function of information known to become available when period t rolls around. At time t, the firm is supposed to have an information set $\Omega_t = \left\{X_t, X_{t-1}, \ldots\right\}$, where X_t is a px1 vector ($p \geq 3$), whose first three elements are P_t, J_t, and W_t. The vector X_t also contains any other variables that the firm finds useful to forecast the process (P,J,W). Of course, at time t, the firm also knows the lagged values k_{t-1} and n_{t-1}. We assume that the vector process X follows the r^{th} order vector autoregressive scheme

$$(1.3) \qquad \delta(L)X_t = V_t^x$$

where $\delta(L) = I - \delta_1 L - \ldots - \delta_r L^r$, where $E\, V_t^x \big| \Omega_{t-1} = 0$, and where the roots of $\det \delta(z) = 0$ are in modulus all greater than $\sqrt{\beta}$. The process for X is assumed to be taken as given by the firm and to be uninfluenced by the firm's decision process.

Substituting the production function (1.2) into the objective (1.1), and using summation by parts and the law of iterated projections to rearrange the expressions in $J_t(k_t - k_{t-1})$ in terms of $(J_t - \beta E_t J_{t+1})k_t$, the objective function can be rewritten

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ P_t d' \begin{bmatrix} k_t \\ n_t \end{bmatrix} - W_t n_t - R_t k_t - \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix}' D \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix} \right\} + J_{-1} k_{-1}$$

where $R_t = J_t - \beta E_t J_{t+1}$. The variable R_t has an interpretation as the rental rate on capital at time t.

The solution of this problem is a linear decision rule of the form

$$\begin{bmatrix} k_t \\ n_t \end{bmatrix} = g_0 + g_1 \begin{bmatrix} k_{t-1} \\ n_{t-1} \end{bmatrix} + h_0 X_t + h_1 X_{t-1} + \ldots h_{r-1} X_{t-r+1}$$

where the g_j's and h_j's are matrices of constants which are complicated functions of the parameters of d, D, and $\delta(L)$. In subsequent sections, we describe quick ways of computing the g_j's and h_j's, and of characterizing their properties.

Example (ii): Market equilibrium with interrelated demand for factors

This is a linear, multiple factor version of the model proposed by Lucas and Prescott [17]. We retain the objective function (1.1), but change the assumption about how the firm forms expectations about the product market price P. In particular, we now take account of the fact that the product price P follows a law of motion that is influenced by the capital accumulation decisions of the aggregate of firms in the industry. First, we assume that R_t and W_t are the first two elements of a (px1) vector Z_t $(p \geq 2)$ which is governed by the r^{th} order autoregressive process

$$(1.4) \qquad \theta(L)Z_t = V_t^z$$

where $\theta(L) = I - \theta_1 L - \dots - \theta_r L^r$, where $E\,V_t^z = E[V_t^z | Z_{t-1}, Z_{t-2}, \dots, U_{t-1}, U_{t-2}, \dots] = 0$, where U is a demand shock to be described below, and where the roots of $\det\theta(z) = 0$ are all greater than $\sqrt{\beta}$ in modulus. Note that P_t is excluded from Z_t. The firm, which is now one of m identical firms, takes the process (1.4) governing Z as given.

The industry demand curve is

$$(1.5) \qquad P_t = A_0 - A_1 Y_t + U_t \quad ; A_0, A_1 > 0$$

where U_t is a random process with mean zero and Y_t is market output. There are m identical firms, so we have

$$(1.6) \qquad Y_t = my_t = m\,d' \begin{bmatrix} k_t \\ n_t \end{bmatrix} = d' \begin{bmatrix} K_t \\ N_t \end{bmatrix}$$

where $K_t = mk_t$ and $N_t = mn_t$. Here K_t and N_t are market-wide stocks of capital and employment, respectively.

We assume that the demand shock U follows the q^{th} order autoregressive specification

$$(1.7) \qquad \xi(L)U_t = V_t^u$$

where $\xi(L) = 1 - \xi_1 L - \dots - \xi_q L^q$, and where $E\, V_t^u = E[V_t^u | Z_{t-1}, Z_{t-2}, \dots, U_{t-1}, U_{t-2}, \dots] = 0$. We also assume that the zeroes of det $\xi(z)$ exceed $\sqrt{\beta}$ in modulus. To complete the model we must specify the firm's beliefs about the motion of K and N, which influence the evolution of the market price P and which the firm has an incentive to forecast. We assume that the representative firm believes that market-wide capital and labor obey the law of motion.

$$(1.8) \qquad \begin{bmatrix} K_t \\ N_t \end{bmatrix} = B_0 + B_1 \begin{bmatrix} K_{t-1} \\ N_{t-1} \end{bmatrix} + G_0 Z_t + \dots + G_{r-1} Z_{t-r+1}$$
$$+ F_0 U_t + \dots + F_{q-1} U_{t-q+1} \quad .$$

Further, the representative firm is assumed to know K_{t-1} and N_{t-1} at t.

Now substituting the demand schedule (1.5) and production function (1.6) into the objective function, the individual firm's problem is to maximize

$$(1.9) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \Big\{ [A_0 - A_1 d' \begin{bmatrix} K_t \\ N_t \end{bmatrix} + U_t] d' \begin{bmatrix} k_t \\ n_t \end{bmatrix} - W_t n_t$$
$$- R_t k_t - \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix}' D \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix} \Big\} + J_{-1} k_{-1}$$

subject to k_{-1}, n_{-1}, K_{-1}, N_{-1} given and known, and the given laws of motion

$$(1.8) \qquad \begin{bmatrix} K_t \\ N_t \end{bmatrix} = B_0 + B_1 \begin{bmatrix} K_{t-1} \\ N_{t-1} \end{bmatrix} + G_0 Z_t + \dots + G_{r-1} Z_{t-r+1}$$
$$+ F_0 U_t + \dots + F_{q-1} U_{t-q+1}$$

$$(1.4) \qquad \theta(L)Z_t = V_t^z$$

$$(1.7) \qquad \xi(L)U_t = V_t^u$$

At time t, the firm's known state variables consist of $k_{t-1}, n_{t-1}, K_{t-1}, N_{t-1}$ and the information variables $\{Z_t, Z_{t-1}, \ldots, U_t, U_{t-1}, \ldots\}$. A solution to this problem is a contingency plan for setting $[k_t, n_t]$ of the linear form

$$(1.10) \qquad \begin{bmatrix} k_t \\ n_t \end{bmatrix} = b_0 + b_1 \begin{bmatrix} k_{t-1} \\ n_{t-1} \end{bmatrix} + b_2 \begin{bmatrix} K_{t-1} \\ N_{t-1} \end{bmatrix} + g_0 Z_t + \ldots + g_{r-1} Z_{t-r+1}$$

$$+ f_0 U_t + \ldots + f_{q-1} U_{t-q+1}.$$

A rational expectations equilibrium is a pair of functions (1.8) and (1.10) such that

$$\begin{bmatrix} K_t \\ N_t \end{bmatrix} = m \begin{bmatrix} k_t \\ n_t \end{bmatrix}$$

identically. This equality between the perceived law of motion for (K,N) and the actual law is readily shown to hold if

$$(1.11) \qquad B_0 = mb_0, \; B_1 = (b_1 + mb_2), \; G_j = mg_j, F_j = m f_j \quad .$$

The restrictions (1.11) guarantee that the individual firm's perception of the law of motion of (K,N) turns out to be accurate, i.e., is implied by optimizing behavior of the individual firms composing the industry. It is to be emphasized that the individual firms are assumed to behave competitively with respect to the market price, and to take as given the process governing the evolution of the market wide stocks of factors, which influence the evolution of the market price.

Notice how the definition of the rational expectations equilibrium builds in: (a) accurate perceptions on the part of firms of the laws of motion for the state processes (K, N, Z, and U) that are beyond their control, (b) optimizing behavior of firms, and (c) market clearing in the output market. However, notice that the model does not analyze "the other side" of the market for inputs, but simply takes the stochastic process Z as given.[1/]

The methods that we describe can readily be used to compute the equilibrium of the model, i.e., the parameters $\{B_0, B_1, G_0, \ldots, G_{r-1}, F_0, \ldots, F_{q-1}\}$ as functions of the underlying parameters d, D, A_0, A_1, $\theta(L)$, and $\delta(L)$. We give an example of such calculations in Section 5.

Example (iii) Exhaustible resource depletion

Epple and Hansen [3] have formulated a model for the purpose of studying the extraction of exhaustible resources. Their model fits a slightly modified version of the "general problem" that we analyze in sections 2 and 3.

Epple and Hansen study a situation in which a vector of resources is being extracted from a single reservoir or mine. They model the exhaustible nature of these resources by positing that marginal exploitation costs increase as a function of the cumulated amount of the resource vector that has been extracted. Resource extraction cost at time t is assumed to be given by

$$\Delta y_t' d + \Delta y_t' S_t + \Delta y_t' D_1 \Delta y_t + \Delta y_t' D_2 (\tfrac{1}{2} \Delta y_t + y_{t-1})$$

where y_t denotes the cumulated amount of the resource extracted as of time t, D_1 and D_2 are positive definite symmetric matrices, d is a column vector, and S is a vector random process representing shocks to the extraction process. The exhaustible nature of the resource is represented by the presence of the quadratic term $\Delta y_t' D_2 (\frac{1}{2} \Delta y_t + y_{t-1})$.

Let P_t be the price vector at time t for the resources. The owner of the resource is assumed to maximize

$$(1.12) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \Big\{P_t' \Delta y_t - \Delta y_t' d - \Delta y_t' S_t - \Delta y_t D_1 \Delta y_t - \Delta y_t' D_2 (\tfrac{1}{2}\Delta y_t + y_{t-1})\Big\}$$

subject to y_{-1} given and subject to the random processes

$$\delta(L)X_t = V_t^x$$

$$\alpha(L)S_t = V_t^s$$

where P_t is the first component of the (pxl) vector X_t, $EV_t^x \mid \Omega_{t-1} = 0$, $EV_t^s \mid \Omega_{t-1} = 0$, and $\Omega_{t-1} = \{X_{t-1}, X_{t-2}, \ldots, S_{t-1}, S_{t-2}, \ldots\}$. The polynomials in the lag operator $\delta(L)$ and $\alpha(L)$ satisfy properties analogous to those specified in example (i). The solution of the owner's maximum problem is a linear contingency plan for setting y_t as a function of $\{y_{t-1}, X_t, X_{t-1}, \ldots, S_t, S_{t-1}, \ldots\}$. In this example, the supplier faces the random process X as a price taker.

The example could be altered to handle the case where the supplier is a monopolist facing a flow demand curve of the linear form

$$(1.13) \qquad P_t = A_0 - A_1 \Delta y_t + U_t, \qquad A_0, A_1 > 0$$

where U_t is a random shock to demand. Alternatively, the example could be modified along the lines of example (ii) to be a model of a competitive industry with a large number of price taking firms with a market flow demand curve such as (1.13).

2. The General Problem

We consider the following setup. Let

y_t be an nx1 vector of variables, typically stocks of things that enter an agent's objective function

h be an nx1 vector of constants

β be a discount factor, with $0 < \beta < 1$

S_{1t} be an nx1 vector of stochastic processes of mean exponential order less than $1/\sqrt{\beta}$.

$S_t = \begin{bmatrix} S_{1t} \\ S_{2t} \end{bmatrix}$ be a (px1) vector of stochastic processes of mean exponential order less than $1/\sqrt{\beta}$. Here $p \geq n$, and S_{1t} is the first n rows of S_t .

H be an nxn positive definite symmetric matrix

$D(L) = D_0 + D_1 L + \ldots + D_m L^m$, D_j an nxn matrix, where D_0 is of full rank, $j = 0,\ldots,m$.

We assume that the px1 vector stochastic process S obeys the r^{th} order autoregression

$$S_t = \delta_1 S_{t-1} + \ldots + \delta_r S_{t-r} + V_t^s$$

or

$$\delta(L) S_t = V_t^s$$

where $\delta(L) = I - \delta_1 L - \ldots - \delta_r L^r$, where δ_j is pxp , and where the zeroes of det $\delta(z)$ are assumed to be greater than $\sqrt{\beta}$ in modulus. This condition on the zeroes of det $\delta(z)$ is equivalent with the condition that S be of mean expotential order less than $(\sqrt{\beta})^{-1}$.

Let E be the mathematical expectations operator. We assume that

$$E\,[V_t^s \mid S_{t-1}, S_{t-2}, \ldots, y_{t-1}, y_{t-2}, \ldots] = 0 \text{ for all } t.$$

$$E\,V_t^s V_t^{s'} = \Sigma_t$$

where Σ_t is a positive semi-definite matrix for all t.[2/] At time t, the agent is assumed to have an information set Ω_t including at least $y_{t-1}, \ldots, y_{t-m}$ and $S_t, S_{t-1}, \ldots, S_{t-r+1}$.

The problem is then to choose a linear decision rule or contingency plan for setting y_t as a function of elements in Ω_t, in order to maximize the objective function

$$\text{(2.1)} \qquad \lim_{N\to\infty} E_0 \sum_{t=0}^{N} \beta^t \left\{(h + S_{1t})' y_t - y_t' H y_t - [D(L)y_t]'[D(L)y_t]\right\}$$

where $E_t(x) = Ex \mid \Omega_t$. The maximization is subject to $y_{-1}, \ldots, y_{-m}$ given and the given law of motion for S_{1t}

$$\text{(2.2)} \qquad \delta(L) \begin{bmatrix} S_{1t} \\ S_{2t} \end{bmatrix} = V_t^s \quad .$$

This problem can be solved as follows by using standard dynamic programming methods for problems with quadratic objectives and linear constraints (e.g., see Bertsekas [1], Kushner [11], or Kwakernaak and Sivan [12]). Define the state vector

$$X_t' = [y_{t-1}', y_{t-2}', \ldots, y_{t-m}', 1, S_t', S_{t-1}', \ldots, S_{t-r+1}'].$$

Define the control vector

$$v_t = D_0 y_t + D_1 y_{t-1} + \ldots + D_m y_{t-m}.$$

The transition equation for the system is then

$$
\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-m+2} \\ y_{t-m+1} \\ 1 \\ S_{t+1} \\ S_t \\ \vdots \\ S_{t-r+3} \\ S_{t-r+2} \end{bmatrix}
=
\begin{bmatrix}
-D_0^{-1}D_1 & -D_0^{-1}D_2 & \cdots & -D_0^{-1}D_{m-1} & -D_0^{-1}D_m & 0 & 0 & 0 & \cdots & 0 & 0 \\
I_n & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 \\
\vdots & & & & & & & & & \vdots & \\
0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 \\
0 & 0 & & I_n & 0 & 0 & 0 & 0 & & 0 & 0 \\
0 & 0 & & 0 & 0 & 1 & 0 & 0 & & 0 & 0 \\
0 & 0 & & 0 & 0 & 0 & \delta_1 & \delta_2 & & \delta_{r-1} & \delta_r \\
0 & 0 & & 0 & 0 & 0 & I_p & 0 & & 0 & 0 \\
\vdots & & & & & & & & & \vdots & \\
0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & I_p & 0
\end{bmatrix}
\begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m+1} \\ y_{t-m} \\ 1 \\ S_t \\ S_{t-1} \\ \vdots \\ S_{t-r+2} \\ S_{t-r+1} \end{bmatrix}
+
\begin{bmatrix} D_0^{-1} \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}
v_t
+
\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ v^{s}_{t+1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}
$$

or $X_{t+1} = A X_t + Bv_t + U_{t+1}$, where A, B and U_{t+1} correspond to the matrices in the line above.

Let $\beta X_t' R X_t$ and $v_t' Q v_t$ be the quadratic forms

$$\begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m} \\ 1 \\ s_t \\ s_{t-1} \\ \vdots \\ s_{t-r+1} \end{bmatrix}' \begin{bmatrix} -H & 0 & \cdots & \frac{1}{2}h & 0 & \frac{1}{2}\phi & \cdots & 0 \\ 0 & 0 & & 0 & 0 & 0 & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{2}h' & 0 & 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & & 0 \\ \frac{1}{2}\phi' & 0 & 0 & 0 & 0 & 0 & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & \cdots\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m} \\ 1 \\ s_t \\ s_{t-1} \\ \vdots \\ s_{t-r+1} \end{bmatrix}$$

and

$$v_t' Q v_t = - v_t' I v_t$$

where $\phi = [I \ 0]$ and is dimensional $n \times p$. Notice that the $(mn) \times (mn)$ submatrix in the upper left hand corner of R is negative semidefinite by virtue of the assumption that H is positive definite

With these definitions of X_t, U_t, v_t, A, B, Q, and R, the problem (2.1) becomes to maximize

(2.3) $$\lim_{N\to\infty} E_0 \sum_{t=0}^{N} \beta^t \left\{ X_t' R X_t + v_t' Q v_t \right\}$$

subject to X_0 given and

(2.4) $$X_{t+1} = A X_t + Bv_t + U_{t+1} \quad .$$

The solution of this control problem is a feedback rule of the linear form

(2.5) $$v_t = -F X_t$$

where

(2.6) $$F = \beta(Q + \beta B' P B)^{-1} B' P A$$

and where P is the "appropriate" solution 3/ of the algebraic matrix Riccati equation

(2.7) $$P = \beta A' P A + R - \beta^2 A' P B (Q + \beta B' P A)^{-1} B' P A \quad .$$

The desired solution of equation (2.7) is obtained by iterations on the matrix Riccati difference equation

(2.8) $$P_{k+1} = \beta A' P_k A + R - \beta^2 A' P_k B (Q + \beta B' P_k B)^{-1} B' P_k A$$

starting from $P_0 = 0$. For our problem, conditions are satisfied that are sufficient to guarantee that iterations on (2.8) converge to the appropriate solution of the algebraic Riccati equation (2.7).4/ The conditions on our problem that are sufficient to guarantee this convergence are (i) that the matrix H is positive definite, and (ii) that the zeroes of det $\delta(z)$ all exceed $\sqrt{\beta}$ in modulus. Furthermore, the conditions on our problem guarantee that the asymptotic closed loop system matrix (A-BF) has all of its eigenvalues less than $1/\sqrt{\beta}$ in modulus.5/

Problems like ours have special features, namely the presence of a distinctive pattern of zeroes in A and B in (2.4), that permit a quicker method of solving problem (2.3).

Let us partition the state X_t' as $X_t' = [X_{1t}', X_{2t}']$ where $X_{1t}' = [y_{t-1}', \ldots, y_{t-m}']$, $X_{2t}' = [1, S_t', \ldots, S_{t-r+1}']$. For applications of the problems that we will study, the dimension of X_1 will generally be much smaller than the dimension of X_2. That is, both the number of variables n being chosen by the "agent", and the number of lags m inherited from the criterion function, will be small relative to the number of variables p in S and the order r of the autoregression for S. Partitioning the rest of the transition equation conformably with $[X_{1t}', X_{2t}']$, we have

$$(2.9) \qquad \begin{bmatrix} X_{1t+1} \\ X_{2t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} v_t + U_{t+1} .$$

This system is in "controllability canonical form", under the assumption that the D_0 is of full rank. This follows from the fact that the pair (A_{11}, B_1) is controllable[6/]. Further, the eigenvalues of A_{22} are all less than $1/\sqrt{\beta}$ by virtue of the assumption that the zeroes of det $\delta(z)$ all exceed $\sqrt{\beta}$ in modulus. Now let us partition P and R conformably with the partition (X_{1t}', X_{2t}') so that

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}' & P_{22} \end{pmatrix} \qquad R = \begin{pmatrix} R_{11} & R_{12} \\ R_{12}' & R_{22} \end{pmatrix} ,$$

where for our problem $R_{22} = 0$. Note again that R_{11} is negative semidefinite by virtue of the assumption that H is positive definite.

By partitioning the algebraic matrix Riccati equation (2.7), it is possible to show that P_{11} is the unique negative definite solution of

$$(2.10) \qquad P_{11} = \beta A_{11}' P_{11} A_{11} + R_{11} - \beta^2 A_{11}' P_{11} B_1 (Q + \beta B_1' P_{11} B_1')^{-1} B_1' P_{11} A_{11}$$

and that P_{11} is the limit as $k \to \infty$ of

$$(2.11) \qquad P_{11,k+1} = \beta A_{11}' P_{11,k} A_{11} + R_{11} - \beta^2 A_{11}' P_{11,k} B_1 (Q + \beta B_1' P_{11,k} B_1)^{-1} B_1' P_{11,k} A_{11}$$

starting from $P_{11,0} = 0$. Also, P_{12} is the stationary point of the difference equation

$$(2.12) \qquad P_{12,k+1} = \beta A_{11}' P_{12,k} A_{22} + R_{12} - \beta^2 A_{11}' P_{11,k} B_1 \left(Q + \beta B_1' P_{11,k} B_1\right)^{-1} B_1' P_{12,k} A_{22}$$

starting from $P_{12,0} = 0$, where the "forcing function" $\left\{P_{11,k}\right\}_{k=0}^{\infty}$ is the solution of (2.11) starting from $P_{11,0} = 0$.[7]

Now partition the control law F in (2.5) conformably with $[X_{1t}', X_{2t}']$, so that

$$v_t = -[F_1 \;\; F_2] \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} .$$

Then it follows from partitioning (2.6) that

$$(2.13) \qquad F_1 = \beta(Q + \beta B_1' P_{11} B_1)^{-1} B_1' P_{11} A_{11}$$

$$(2.14) \qquad F_2 = \beta(Q + \beta B_1' P_{11} B_1)^{-1} B_1' P_{12} A_{22}$$

The piece $F_1 X_{1t}$ of the control law is sometimes called the "feedback" part of the control law, while $F_2 X_{2t}$ is called the "feedforward" part.

From the matrix Riccati equation (2.10) for P_{11} and from the formula (2.13) for F_1, it follows that the feedback matrix F_1 depends only on the matrices A_{11}, B_1, and R_{11}, and is independent of the parameters characterizing the random process S. Recall that the feedback matrix F_1 gives the dependence of the control law for v_t on the initial state variables $X_{1t}' = [y_{t-1}', y_{t-2}', \ldots, y_{t-m}']$.

From inspection of the Riccati equation (2.10) for P_{11} and the formula (2.13) for F_1, it follows that the parameters of F_1 are exactly those that come from solving the following smaller-dimensional, nonstochastic optimization problem:

$$\text{maximize} \sum_{t=0}^{\infty} \beta^t (X_{1t}' R_{11} X_{1t} + v_t' Q v_t)$$

subject to

$$X_{1t+1} = A_{11} X_{1t} + B_1 v_t$$

$$X_{10} \text{ given} \quad .$$

The solution of this problem is the feedback law

$$v_t = -F_1 X_{1t}$$

where F_1 is given by equation (2.13). Further, notice that $(A_{11}-B_1F_1)$ is just the upper left square of (A-BF), which has the block triangular form

$$\begin{pmatrix} A_{11}-B_1F_1 & A_{12}-B_1F_2 \\ 0 & A_{22} \end{pmatrix}$$

It follows from the earlier result that the eigenvalues of (A-BF) are bounded in modulus by $\beta^{-1/2}$ that the eigenvalues of $(A_{11}-B_1F_1)$ are bounded in modulus by $\beta^{-1/2}$. This bound on the modulus of the eigenvalues of $(A_{11}-B_1F_1)$ is an important feature of our problem, which we propose to exploit in devising an alternative solution procedure in the following section of this paper. In particular, it is this feature of our problem that verifies that the particular solution that we choose to the Euler equations described in the following section is the optimizing choice.[8]

Now it happens that given F_1, the parameters of F_2 can be found directly without the need to use (2.14) and without the need to iterate on (2.12). This is true because of a certain kind of symmetry between the feedback and feedforward parts of the optimal control law. Below, we note this symmetry for our problem and show how it can be combined with the Wiener-Kolmogorov theory of linear least squares prediction to compute F_2 directly from knowledge of F_1 and $\delta(L)$. There will be computational and other advantages from pursuing this strategy.

3. Solutions Using Wiener-Kolmogorov Formulas

We return to our problem (2.1). We propose to solve the problem by using the certainty equivalence principle. That is, we shall first solve a version of the problem assuming that $\{S_{1t}\}$ is a known sequence, rather than a stochastic process. We derive an expression for the decision rule in which y_t depends linearly on lagged y's and actual future $S_{1t+j}, j \geq 0$. By the certainty equivalence principle, the correct rule under uncertainty can be derived from this rule by replacing S_{1t+j} for $j \geq 1$ with $E_t S_{1t+j}$, the linear least squares forecast of S_{1t+j} based on information available at time t. We use earlier results of Hansen and Sargent [7], which are based on the classic Wiener-Kolmogorov prediction theory, to derive a convenient operational expression for the part of the decision rule that implicitly reflects the $E_t S_{1t+j}$'s. This procedure leads to the optimal linear decision rule.

We first solve the certainty version of our problem, maximize

$$(3.1) \qquad J = \lim_{N\to\infty} \sum_{t=0}^{N} \beta^t \left\{(h+S_{1t})'y_t - y_t' Hy_t - [D(L)y_t]'[D(L)y_t]\right\}$$

where $\{S_{1t}\}_{t=0}^{\infty}$ is regarded as a known sequence, and where the maximization is over sequences $\{y_t\}_{t=0}^{\infty}$. The initial conditions $y_{-1},\ldots,y_{-m}$ are all given. To solve the problem, we fix $N \gg m$, and consider first the N-period problem (3.1) with N fixed. We shall obtain a set of first order necessary conditions for a maximum of (3.1) with N fixed.

Consider first the term

$$I = \sum_{t=0}^{N} \beta^t [D(L)y_t]'[D(L)y_t]$$

$$= \sum_{t=0}^{N} \beta^t [y_t'D_0' + y_{t-1}'D_1' + \ldots + y_{t-m}'D_m'] \; [D_0 y_t + \ldots + D_m y_{t-m}] \; .$$

Differentiating I with respect to y_t for $t = 0,\ldots,N\text{-}m$ gives

$$
\begin{aligned}
\frac{\partial I}{\partial y_t} &= \beta^t D_0' D(L) y_t + \beta^{t+1} D_1' D(L) y_{t+1} + \ldots + \beta^{t+m} D_m' D(L) y_{t+m} \\
&\quad + \beta^t \left\{ [D(L)y_t]' D_0 \right\}' + \beta^{t+1} \left\{ [D(L)y_{t+1}]' D_1 \right\}' \\
&\quad + \ldots + \beta^{t+m} \left\{ [D(L)y_{t+m}]' D_m \right\}' \\
&= 2\beta^t [D_0' D(L) + \beta L^{-1} D_1' D(L) + \ldots + \beta^m L^{-m} D_m' D(L)] y_t \\
&= 2\beta^t \, D(\beta L^{-1})' D(L) y_t \quad .
\end{aligned} \tag{3.2}
$$

Therefore, to maximize (3.1) with N taken fixed, we have the Euler equations

$$
\frac{\partial J}{\partial y_t} = \beta^t [(h + S_{1t}) - 2Hy_t - 2D(\beta L^{-1})' D(L) y_t] = 0 \tag{3.3}
$$

$$
t = 0, 1, \ldots, N\text{-}m
$$

For the infinite time problem obtained by driving N to infinity, the Euler equations (3.3) are also first order necessary conditions. However, as there exists more than one solution of the Euler equations (3.3) that satisfies the mxn initial conditions $y_{-1}, \ldots, y_{-m}$, we need additional conditions to pick out the unique optimum path of y_t. For the certainty version of our problem, it can be proved that as $N \to \infty$, the optimum of criterion function (3.1) is bounded below.[9/] This means that the optimum path for y_t satisfies

$$
\sum_{t=0}^{\infty} \beta^t y_t' H y_t < +\infty \tag{3.4}
$$

This condition is shown in the Appendix A uniquely to determine the solution of the Euler equations that our procedure selects. Equivalently, our procedure is known to be correct because it selects the unique solution of the Euler equations that gives rise to a closed loop system with zeroes of its characteristic polynomial that are bounded in modulus by $\beta^{-1/2}$. As indicated in Section 2, this is known to be a property of the optimal closed loop system.

Write the Euler equation as

(3.5) $$[H+D(\beta L^{-1})'D(L)]y_t = \frac{1}{2}(h+S_{1t}).$$

The Euler equations (3.5) can be solved subject to (3.4) and n·m initial conditions, $y_{-1}, \ldots, y_{-m}$, by the following procedure. First, we note that the roots of det $[H+D(\beta z^{-1})'D(z)] = 0$ come in pairs: if z_0 is a root, so is βz_0^{-1}. It is, in general, possible to factor the matrix polynomial $H + D(\beta z^{-1})'D(z)$ so that

(3.6) $$H + D(\beta z^{-1})'D(z) = C(\beta z^{-1})'C(z)$$

where $C(z)$ is an m^{th} order, nxn matrix polynomial in nonnegative powers of z, $C(z) = C_0 + C_1 z + \ldots + C_m z^m$, and where all of the roots of det $C(z) = 0$ in modulus are greater than $\sqrt{\beta}$. For each root z_0 of det $C(z) = 0$, there is a root βz_0^{-1} of det $C(\beta z^{-1}) = 0$. The roots of det $C(\beta z^{-1}) = 0$ in modulus are all less than $\sqrt{\beta}$. The factorization (3.6) is unique up to premultiplication of $C(L)$ by an orthogonal matrix. These assertions about (3.6) are proved in Appendix B.

Using the factorization (3.6), we can write the Euler equations (3.5) as

$$C(\beta L^{-1})'C(L)y_t = \frac{1}{2}(h+S_{1t}).$$

The solution of this equation that satisfies condition (3.4) is then

$$C(L)y_t = \frac{1}{2}[C(\beta L^{-1})']^{-1}(h + S_{1t}) \tag{3.7}$$

or

$$C_0 y_t + C_1 y_{t-1} + \ldots + C_m y_{t-m} = \tag{3.8}$$

$$\frac{1}{2}(C_0' + C_1'\beta L^{-1} + \ldots + C_m'\beta^m L^{-m})^{-1}(h + S_{1t}).$$

As shown in Appendix A, condition (3.4) and the initial conditions impel us to solve the "stable roots backwards and the unstable roots forwards". Now premultiplying both sides of (3.8) by C_0^{-1} gives

$$y_t + C_0^{-1}C_1 y_{t-1} + \ldots + C_0^{-1}C_m y_{t-m} = \tag{3.9}$$

$$\frac{1}{2}(C_0'C_0 + C_1'C_0\beta L^{-1} + \ldots + C_m'C_0\beta^m L^{-m})^{-1}(h + S_{1t}) \quad .$$

From section 2 we have a quick and feasible method of obtaining F_1', which together with D(L) directly gives the "feedback" polynomial in (3.9), namely $[I + C_0^{-1}C_1 L + C_0^{-1}C_2 L^2 + \ldots + C_0^{-1}C_m L^m]$.
Given this polynomial, we shall now describe a method for obtaining a tractable expression for the "feedforward" part of the solution in (3.9), namely

$$\frac{1}{2}(C_0'C_0 + C_1'C_0\beta L^{-1} + \ldots + C_m'C_0\beta^m L^{-m})^{-1}(h + S_{1t}) \; .$$

First, by multiplying the polynomials in L, it is established that $C(\beta L^{-1})'C(L) = \sum_{j=-m}^{m} \tilde{C}_j L^j$ where

$$\tilde{C}_0 = C_0'C_0 + \beta C_1'C_1 + \ldots + \beta^m C_m'C_m$$

$$\tilde{C}_1 = C_0'C_1 + \beta C_1'C_2 + \ldots + \beta^{m-1} C_{m-1}'C_m$$

(3.10) $$\tilde{C}_2 = C_0'C_2 + \beta C_1'C_3 + \ldots + \beta^{m-2} C_{m-2}'C_m$$

$$\vdots$$

$$\tilde{C}_m = C_0'C_m$$

$$\tilde{C}_{-j} = \beta^j \tilde{C}_j' \text{ for } j = 1,\ldots,m .$$

Similarly, where $D(\beta L^{-1})'D(L) = \sum_{j=m}^{m} \tilde{D}_j L^j$, it is established that

$$\tilde{D}_0 = D_0'D_0 + \beta D_1'D_1 + \ldots + \beta^m D_m'D_m$$

(3.11) $$\vdots$$

$$\tilde{D}_m = D_0'D_m$$

$$\tilde{D}_{-j} = \beta^j \tilde{D}_j', \; j = 1,\ldots,m .$$

Therefore, given D(L), it is evident from (3.10), (3.11), and (3.6) that

(3.12) $$C_0' C_m = D_0'D_m \quad .$$

Next, since $C_0^{-1}C_1,\ldots,C_0^{-1}C_m$ are all known from the feedback polynomial in (3.9), we can calculate $C_0'C_0$ using (3.6) from

$$C_0'C_0 = C_0'C_m(C_0^{-1}C_m)^{-1}$$

or

(3.13) $$C_0'C_0 = (D_0'D_m)(C_0^{-1}C_m)^{-1} .$$

Equation (3.13) expresses $C_0'C_0$ in terms of known parameters. Given $C_0'C_0$, we can obtain $C_j'C_0$ from

$$(3.14) \qquad C_j'C_0 = (C_0^{-1}C_j)'C_0'C_0, \; j = 1, \ldots, m.$$

So equations (3.13) and (3.14) together with the known values of $D_0'D_m$ and $C_0^{-1}C_j$ determine all the matrices that appear on the right side of (3.9):

$$M(\beta z^{-1}) = C_0'C_0 + C_1'C_0\beta z^{-1} + \ldots + C_m'C_0\beta^m z^{-m}.$$

In order to compute the equilibrium decision rule (3.9), we shall need a convenient algorithm for computing the inverse of $M(z)$, since the inverse of $M(\beta L^{-1})$ appears on the right side of (3.9). We now describe convenient formulas for computing $M(z)^{-1}$ via the identity $M(z)^{-1} = \text{adj } M(z)/\det M(z)$. We proceed by describing algorithms for computing both adj $M(z)$ and det $M(z)$. These calculations ultimately lead to equation (3.18) below. The intervening calculations are useful technical details which can be skipped on first reading.

Our procedure is an adaptation of one proposed by Emre and Hüseyin [2]. We begin by evaluating det $M(z)$. To accomplish this task we note that

$$(3.15) \qquad \frac{\partial \det M(z)}{\partial z} = \text{trace } [\frac{\partial \det M(z)}{\partial M(z)} \; \frac{\partial M(z)'}{\partial z}]$$

$$= \det M(z) \text{ trace } [\frac{\partial M(z)}{\partial z} M(z)^{-1}]$$

Let $\tilde{M}(z) = \left(\frac{\partial M(z)}{\partial z}\right) M(z)^{-1}$ and write its Taylor series expansion about zero as

$$\tilde{M}(z) = \tilde{M}_0 + \tilde{M}_1 z + \tilde{M}_2 z^2 + \ldots \; .$$

Now

$$\tilde{M}(z)M(z) = \frac{\partial M(z)}{\partial z}$$

$$= C_1'C_0 + 2C_2'C_0 z + \ldots + mC_m'C_0 z^{m-1}.$$

Equating coefficients in the Taylor series we know that

$$\tilde{M}_0 C_0' C_0 = C_1' C_0 \quad \text{or} \quad \tilde{M}_0 = C_1' C_0'^{-1}$$

$$\tilde{M}_1 C_0' C_0 + \tilde{M}_0 C_1' C_0 = 2C_2' C_0 \quad \text{or} \quad \tilde{M}_1 = 2C_2' C_0'^{-1} - \tilde{M}_0 C_1' C_0'^{-1}$$

$$\vdots$$

$$\tilde{M}_j C_0' C_0 + \ldots + \tilde{M}_0 C_j' C_0 = (j+1)\ C_{j+1}' C_0 \quad \text{or}$$

$$\tilde{M}_j = (j+1)\ C_{j+1}' C_0'^{-1} - \tilde{M}_{j-1} C_1' C_0'^{-1} - \ldots - \tilde{M}_0 C_j' C_0'^{-1}$$

where $C_j = 0$ for $j > m$.

This provides us with recursive formulas for the $\tilde{M}_j$'s. Let $k = m \cdot n$ and write

$$\det M(z) = d_0 + d_1 z + \ldots + d_k z^k .$$

Differentiating with respect to z we obtain

$$\frac{\partial \det M(z)}{\partial z} = d_1 + 2d_2 z + \ldots + kd_k z^{k-1} .$$

We can rewrite equation (3.15) as

$$\frac{\partial \det M(z)}{\partial z} = \det M(z)\,[\text{trace } \tilde{M}_0 + \text{trace } \tilde{M}_1 z + \text{trace } \tilde{M}_2 z^2 + \ldots] .$$

Again we equate coefficients to obtain

$$d_1 = d_0 \text{ trace } \tilde{M}_0$$

$$2d_2 = d_0 \text{ trace } \tilde{M}_1 + d_1 \text{ trace } \tilde{M}_0 \quad \text{or} \quad d_2 = \tfrac{1}{2}(d_0 \text{ trace } \tilde{M}_1 + d_1 \text{ trace } \tilde{M}_0)$$

$$\vdots$$

$$kd_k = d_0 \text{ trace } \tilde{M}_{k-1} + \ldots + d_{k-1} \text{ trace } \tilde{M}_0 \quad \text{or}$$

$$d_k = \frac{1}{k}(d_0 \text{ trace } \tilde{M}_{k-1} + \ldots + d_{k-1} \text{ trace } \tilde{M}_0)$$

Noting that $d_0 = \det(C_0'C_0)$ we have now derived recursive formulas for the d_j's and thus the polynomial coefficients for $\det M(z)$. Using these coefficients and a numerical factorization algorithm we can express

$$\det M(z) = d_k(z-z_1)(z-z_2)\ldots(z-z_k) .$$

Where $z_1,\ldots,z_k$ are the roots of $\det M(z)$. These roots are greater than $\sqrt{\beta}$ in magnitude.

In order to proceed to the second step of the inversion of $M(z)$, we shall write

$$M(z)^{-1} = \frac{\operatorname{adj} M(z)}{d_k(z-z_1)(z-z_2)\ldots(z-z_k)} \tag{3.16}$$

where $\operatorname{adj} M(z)$ is the adjoint of $M(z)$. Thus our second step involves deriving a formula for $\operatorname{adj} M(z)$. The Taylor series expansion about zero for $\operatorname{adj} M(z)$ can be written

$$\operatorname{adj} M(z) = M_0^* + M_1^* z + \ldots + M_{k-m}^* z^{k-m}.$$

Notice that $[\operatorname{adj} M(z)]M(z) = [\det M(z)]I$; equating coefficients of the Taylor series expansion of both sides of this equation gives

$$M_0^* C_0' C_0 = d_0 I \qquad \text{or} \quad M_0^* = d_0(C_0'C_0)^{-1}$$

$$M_0^* C_1' C_0 + M_1^* C_0' C_0 = d_1 I \quad \text{or} \quad M_1^* = d_1(C_0'C_0)^{-1} - M_0^* C_1' C_0'^{-1}$$

$$\vdots$$

$$M_0^* C_{k-m}' C_0 + \ldots + M_{k-m}^* C_0' C_0 = d_{k-m} I$$

or

$$M_{k-m}^* = d_{k-m}(C_0'C_0)^{-1} - M_0^* C_{k-m}' C_0'^{-1} \ldots - M_{k-m-1}^* C_1' C_0'^{-1}.$$

This provides us with recursive formulas for the M_j^*'s.

The third step in our inversion formula for $M(z)$ amounts to expanding (3.16) by matrix partial fractions to obtain

$$M(z)^{-1} = \frac{N_1}{(z-z_1)} + \ldots + \frac{N_k}{(z-z_k)} \tag{3.17}$$

$$\text{where} \quad N_j = \frac{1}{d_k \prod\limits_{\substack{i \neq j \\ 1 \leq i \leq k}} (z_j - z_i)} [M_0^* + M_1^* z_j + \ldots + M_{k-m}^* (z_j)^{k-m}].$$

Equation (3.17) can be rewritten

$$M(z)^{-1} = \frac{-\frac{1}{z_1} N_1}{(1-\frac{1}{z_1} z)} + \ldots + \frac{-\frac{1}{z_k} N_k}{(1-\frac{1}{z_k} z)} \tag{3.18}$$

Finally, substituting βz^{-1} for z gives

$$M(\beta z^{-1})^{-1} = (C_0'C_0 + C_1'C_0\beta z^{-1} + \ldots + C_m'C_0\beta^m z^{-m})^{-1} \tag{3.19}$$

$$= \frac{-\frac{1}{z_1} N_1}{1-\frac{1}{z_1}\beta z^{-1}} + \ldots + \frac{-\frac{1}{z_k} N_k}{1-\frac{1}{z_k}\beta z^{-1}}$$

Applying equation (3.19) to the right side of (3.9) gives the decision rule for y_t:

$$y_t + C_0^{-1} C_1 y_{t-1} + \ldots + C_0^{-1} C_m y_{t-m} = \tag{3.20}$$

$$\frac{1}{2} \sum_{j=1}^{k} \left[\frac{-\lambda_j N_j}{1-\lambda_j \beta L^{-1}} \right] (h + S_{1t})$$

where $\lambda_j = \frac{1}{z_j}$.

Using the expansion $[1 - (\lambda_j)\beta L^{-1}]^{-1} = \sum_{i=0}^{\infty} (\lambda_j\beta)^i L^{-i}$, (3.20) can be written

(3.21)
$$y_t = -(C_0^{-1}C_1 y_{t-1} + \dots + C_0^{-1}C_m y_{t-m})$$
$$- \frac{1}{2}\sum_{j=1}^{k} \lambda_j N_j \left[\sum_{i=0}^{\infty} (\lambda_j\beta)^i (S_{1t+i} + h)\right]$$

Equation (3.20) or (3.21) expresses the optimal choice of y_t as a function of m lagged y's and the sum of k geometric sums of all future values of the vector sequence S_1.

It is now a simple step to add uncertainty. Where S is a random process obeying the assumptions we have imposed above, the optimal rule is obtained by replacing S_{1t+i} with $E_t S_{1t+i}$ in (3.21):

(3.22)
$$y_t = -(C_0^{-1}C_1 y_{t-1} + \dots + C_0^{-1}C_m y_{t-m})$$
$$- \frac{1}{2}\sum_{j=1}^{k} \lambda_j N_j \left[\sum_{i=0}^{\infty} (\lambda_j\beta)^i (E_t S_{1t+i} + h)\right] .$$

By using a formula of Hansen and Sargent [7], the geometric sum in expected S_{1t+i}'s can be written

(3.23)
$$\sum_{i=0}^{\infty} (\lambda_j\beta)^i E_t S_{1t+i} = \phi\,\delta(\lambda_j\beta)^{-1} \left\{I + \sum_{s=1}^{r-1} \left\{\sum_{i=s+1}^{r} (\lambda_j\beta)^{i-s}\delta_i\right\} L^s\right\} S_t,$$

where ϕ is an n x p matrix of the form [I 0]. The classic Wiener-Kolmogorov prediction formulas are embedded in (3.23), as described by Hansen and Sargent [7]. Substituting (3.23) into (3.22), we obtain the decision rule

(3.24)
$$y_t = -(C_0^{-1} C_1 y_{t-1} + \dots + C_0^{-1} C_m y_{t-m})$$
$$- \frac{1}{2} \sum_{j=1}^{k} \lambda_j N_j \phi\, \delta(\lambda_j \beta)^{-1} \left\{ I + \sum_{s=1}^{r-1} \left[\sum_{i=s+1}^{r} (\lambda_j \beta)^{i-s} \delta_i \right] L^s \right\} S_t$$
$$- \frac{1}{2} \sum_{j=1}^{k} \lambda_j N_j \left(\frac{1}{1-\lambda_j \beta}\right) h \quad .$$

Equation (3.24) expresses the optimal choice of y_t as a function of m lagged y's and current and (r-1) lagged values of S. The "state" variables thus match up with the setup of section 2.

In Section 5, we shall exhibit speeds of calculating a particular numerical example of a Lucas-Prescott equilibrium of investment uncertainty, using both the method of Section 2 and the method leading up to (3.24). This will give the reader some sense of how much quicker the method of this section can be than the earlier one. Before proceeding to this example, in the next section we describe a modification of the present procedure which differs in that it factors the matrix $H + D(\beta L^{-1})'D(L)$ by a different method.

4. Another Solution Procedure

The previous two sections have indicated two different but related methods for solving our general problem. The first method involved casting the problem in the form of a "stochastic linear optimal regulator" problem, and solving it by iterating on the matrix Riccati difference equation. This approach in effect solved the "optimization" and "prediction" pieces of the problem jointly. The second method explicitly separated the optimization and the prediction problems, used the recursive method to factor the spectral density-like matrix $[H+D(\beta L^{-1})D(L)]$ involved in the "optimization" piece of the problem, but used analytic, nonrecursive formulas to solve the "prediction" aspects of the problem.

A third procedure is also available in principle, and it is practical in sufficiently small systems (n and m should be small). The method involves using the procedure described by Rozanov [21] to factor $[H+D(\beta L^{-1})'D(L)]$.10/ By using this procedure, the requirement for nonanalytic procedures (i.e., numerical or recursive procedures) can be reduced to the minimum extent possible, namely, to the need to find the roots of several univariate polynomials. In general, the procedure can be described as follows. The matrix characteristic polynomial for the Euler equations can be represented, as in Appendix B, as

$$(4.1)\qquad H+D(\beta L^{-1})'D(L) = H+\bar{D}(\sqrt{\beta}\,L^{-1})'\,\bar{D}\left(\frac{1}{\sqrt{\beta}}L\right)$$

where $\bar{D}_j = D_j(\sqrt{\beta})^j$. We use Rozanov's procedure to factor the "spectral density" matrix

$$(4.2)\qquad H+\bar{D}(z^{-1})'\,\bar{D}(z) = G(z^{-1})'\,G(z)$$

where $G_j = \sum_{j=0}^{m} G_j z^j$, and the roots of det $G(z) = 0$ are all outside the unit circle. As in Appendix B, the Euler equations for the certainty version of our problem can be written as

$$(4.3) \qquad G(\sqrt{\beta}L^{-1})' \, G(\tfrac{1}{\sqrt{\beta}}L)y_t = \tfrac{1}{2}(h + S_t)$$

or

$$(4.4) \qquad C(\beta L^{-1})' \, C(L)y_t = \tfrac{1}{2}(h + S_{1t})$$

where $C_j = \left(\frac{1}{\sqrt{\beta}}\right)^j G_j$. Once C(L) has been obtained, the solution for the feedback rule can be obtained exactly as described in Section 3.

The advantage of using the method described by Rozanov over the methods of sections 2 and 3 is that Rozanov's delivers closed form expressions (or "nearly" closed form expressions, depending on the size of m and n). The disadvantage of Rozanov's method vis a vis using the recursive method of section 3 is that the algebraic calculations required for Rozanov's method are tedious.

In Appendix C, we report explicit closed term formulas for factoring $H + \bar{D}(z)'\bar{D}(z)$ for the case in which $n = 2$, $m = 1$. These formulas were derived by following the instructions provided by Rozanov [21]. When combined with formulas (3.15) -(3.24), the formulas in Appendix C give a completely closed form expression for the decision rule in the $n = 2$, $m = 1$ case.

5. An Illustration

We illustrate these computational methods by computing the equilibrium of a multiple factor version of the Lucas-Prescott model, which was the second example in Section 1. Recall that the firm is assumed to maximize

$$(5.1)\qquad E_0 \sum_{t=0}^{\infty} \beta^t \{P_t d' \begin{bmatrix} k_t \\ n_t \end{bmatrix} - W_t n_t - R_t k_t - \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix}' D \begin{bmatrix} \Delta k_t \\ \Delta n_t \end{bmatrix}\}$$

subject to

$$(5.2)\qquad P_t = A_0 - A_1 Y_t + U_t$$

$$(5.3)\qquad \begin{bmatrix} K \\ N \end{bmatrix}_t = B_0 + B_1 \begin{bmatrix} K_{t-1} \\ N_{t-1} \end{bmatrix} + G_0 Z_t + \dots + G_{r-1} Z_{t-r+1} + F_0 U_t + \dots + F_{q-1} U_{t-q+1}$$

$$(5.4)\qquad \theta(L) Z_t = V_t^z$$

$$(5.5)\qquad \xi(L) U_t = V_t^u$$

where the rentals R_t and W_t are the first two elements of the $(p \times 1)$ vector process Z_t. At time t, the firm is assumed to know the state variables $\{k_{t-1}, n_{t-1}, K_{t-1}, N_{t-1}\}$ and the information variables $\{Z_t, Z_{t-1}, \dots, U_t, U_{t-1}, \dots\}$. The firm knows the parameters of the laws of motion for $(K, N)'$, Z, and U, and also the parameters of the demand schedule.

To compute the equilibrium law of motion for $(K_t, N_t) \equiv (mk_t, mn_t)$, we follow Lucas and Prescott [17] or Sargent [23] and solve the following social planning problem: to maximize[11/]

$$(5.6) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \{A_0 m d' \begin{bmatrix} k_t \\ n_t \end{bmatrix} - \tfrac{1}{2} A_1 m^2 \begin{bmatrix} k_t \\ n_t \end{bmatrix}' d d' \begin{bmatrix} k_t \\ n_t \end{bmatrix}$$

$$+ md' \begin{bmatrix} k_t \\ n_t \end{bmatrix} U_t - m \begin{bmatrix} R_t \\ W_t \end{bmatrix}' \begin{bmatrix} k_t \\ n_t \end{bmatrix}$$

$$- m \begin{pmatrix} k_t - k_{t-1} \\ n_t - n_{t-1} \end{pmatrix}' D \begin{pmatrix} k_t - k_{t-1} \\ n_t - n_{t-1} \end{pmatrix} \}.$$

In (5.6), the maximization is over contingency plans for $[k_t, n_t]$ given the laws of motion (5.4) and (5.5), and given $[k_{t-1}, n_{t-1}, Z_t, Z_{t-1}, \ldots, U_t, U_{t-1}, \ldots]$. Once the contingency plan for the representative firm's stocks $[k_t, n_t]$ that maximizes (5.6) has been obtained, the competitive equilibrium for $[K_t, N_t]$ can be obtained by multiplying by m, that is, by using $[K_t, N_t] = [mk_t, mn_t]$. It should be noted that the contingency plan for $[k_t, n_t]$ that maximizes the social planning criterion (5.6) is not the optimum contingency plan (1.10) of the representative firm of section 1, but is simply m^{-1} times the equilibrium law of motion (1.8) for $[K,N]$. That the competitive equilibrium described in section 1 implicitly maximizes (5.6) can be verified directly by using an argument analogous to that of Sargent [22].

Using each of our three methods, we have calculated a rational expectations equilibrium by maximizing (5.6). For the Z process (5.4) we assumed

$$\begin{pmatrix} R_t \\ W_t \end{pmatrix} = \begin{pmatrix} .6 & .2 \\ .7 & -.1 \end{pmatrix} \begin{pmatrix} R_{t-1} \\ W_{t-1} \end{pmatrix} + \begin{pmatrix} -.2 & .3 \\ .1 & -.1 \end{pmatrix} \begin{pmatrix} R_{t-2} \\ W_{t-2} \end{pmatrix}$$

$$+ \begin{pmatrix} -.1 & -.4 \\ .3 & .2 \end{pmatrix} \begin{pmatrix} R_{t-3} \\ W_{t-3} \end{pmatrix} + \begin{pmatrix} .1 & .0 \\ -.1 & .2 \end{pmatrix} \begin{pmatrix} R_{t-4} \\ W_{t-4} \end{pmatrix} + V_t^z$$

For (5.5), we assumed for simplicity that $U_t \equiv 0$, so that demand shocks are suppressed. (It would be very cheap to include demand shocks with our second and third computational methods, somewhat more expensive with the first method.) We assumed that $A_1 = .00005$, $m = 1000$, $\beta = .9$, and $d' = (.25, .75)$. We chose D to obey

$$mD = \begin{pmatrix} 2 & 1 \\ 1 & 1.5 \end{pmatrix}.$$

To match the objective function (5.6) of the social planning problem with the objective function (2.1) of the general optimization problem described in section 2 above, we set

$$H = \tfrac{1}{2}A_1 m^2 dd',$$

$$h = A_0 md,$$

$$[D(L)y_t]'[D(L)y_t] = \begin{bmatrix} k_t - k_{t-1} \\ n_t - n_{t-1} \end{bmatrix}'[mD]\begin{bmatrix} k_t - k_{t-1} \\ n_t - n_{t-1} \end{bmatrix},$$

and

$$S_t \equiv S_{1t} = -m\begin{bmatrix} R_t \\ W_t \end{bmatrix}.$$

Also, we set $A_0 = 0$, which amounts to setting constant terms in the equilibrium [K,N] process to zero. The resulting equilibrium should then be thought of as describing variables measured in deviations from their means.[12]

We calculated the equilibrium three ways: with the method of section 2 which involves iterating on the full matrix Riccati difference equation (2.8), with the "short" method of section 3, and with the "shorter" method of section 4. The equilibrium can be written as

$$(5.7)\qquad \begin{pmatrix} k_t \\ n_t \end{pmatrix} = \begin{pmatrix} 1.1021 & .3064 \\ -.3404 & -.0213 \end{pmatrix} \begin{pmatrix} k_{t-1} \\ n_{t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} -.4977 & .0627 \\ .1747 & -.0542 \end{pmatrix} \begin{pmatrix} mR_t \\ mW_t \end{pmatrix} + \begin{pmatrix} .0815 & .0378 \\ -.0275 & -.0122 \end{pmatrix} \begin{pmatrix} mR_{t-1} \\ mW_{t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} -.0139 & .1960 \\ .0053 & -.0661 \end{pmatrix} \begin{pmatrix} mR_{t-2} \\ mW_{t-2} \end{pmatrix} + \begin{pmatrix} -.0493 & .0094 \\ .0167 & -.0036 \end{pmatrix} \begin{pmatrix} mR_{t-3} \\ mW_{t-3} \end{pmatrix}$$

To express the equilibrium in terms of $[K_t, N_t]$, simply multiply both sides of (5.7) by $m = 1000$.

To implement the first two computational methods, which are iterative, a convergence criterion had to be adopted. We used the following convergence criterion. We calculated successive iterates on the feedback law, namely,

$$F_k = \beta(Q+\beta B'P_kB)^{-1}B'P_kA,$$

where iterations on the matrix Riccati difference equation (2.7) were started from $P_0 = 0$. Then we computed the norm defined as the maximum absolute value over elements of $(F_{k+1}-F_k)$. For convergence, we insisted that this norm had to be less than 10^{-5}. For the $[k_t, n_t]$ law that maximized (5.6), all three methods gave identical answers to at least five digits, as expected. (The results for the third method involve no iteration and are exact.)

Table 1 gives the time taken for each method in central processor time on the Cyber 172 at the University of Minnesota. Generally, we would expect the relative speed advantage of the shorter methods in calculating the equilibrium to increase the closer is β to unity and the larger is the dimensionality of the S process both in terms of the number of lags in its autoregression, and the number of variables in S. For our example, since central processor time costs about eight cents per second, one evaluation of the equilibrium costs about half a penny by the short methods, about 25 cents by the full Riccati method. The relative costliness of these computational procedures clearly will vary from problem to problem.

Table 1

Central Processor Time to Calculate Rational Expectations Equilibrium, in Seconds*

Method	Time
Full Riccati (section 2)	3.247
Short Riccati (section 3)	.075
Spectral Factorization (section 4)	.052

*On Cyber 172 Computer, University of Minnesota

6. A Possible Identification Problem

Let us represent the decision rule (3.24), using (3.8) in the compact form

$$C(L)y_t = E_t\left\{\frac{1}{2}[C(\beta L^{-1})']^{-1}\ (h + S_{1t})\right\} \tag{6.1}$$

where $\delta(L)S_t = v_t^s$. Here it is understood that E_t stands for the expectation or linear projection conditioned on $\left\{S_t, S_{t-1}, \ldots\right\}$.

Suppose the econometrician sees enough of the (y,S) process to permit him to estimate the parameters of $C(L)$ and h. Let us pose the following question. Is it possible to work backwards to obtain a unique H and $D(L)$ such that

$$H + D(\beta L^{-1})'\ D(L) = C(\beta L^{-1})'\ C(L)?$$

In other words, can we identify the criterion function parameters H and $D(L)$ from the decision rule parameters $C(L)$? First of all, there is a relatively trivial sense in which the answer to this question is no. Both $C(L)$ and $D(L)$ are identified only up to a premultiplication by a unitary matrix. From the standpoint of criterion function identification, this problem is not particularly interesting because premultiplication of $D(L)$ by an orthogonal matrix does not effect the term

$$[D(L)y_t]'\ [D(L)y_t]$$

that enters into the criterion function. In other words if Λ is an orthogonal matrix conformable with $D(L)$, then

$$[\Lambda D(L)y_t]'\ [\Lambda D(L)y_t]$$

$$= [D(L)y_t]\Lambda'\Lambda[D(L)y_t]$$

$$= [D(L)y_t]'\ [D(L)y_t]\quad .$$

This suggests that all we should really care about is identification of D(L) up to a premultiplication by an orthogonal matrix since elements in this class of D(L)'s all give rise to the same criterion function.

It turns out that there is another sense in which the criterion function parameters cannot be identified from the decision rule parameters. Using the procedure suggested in section 4 and Appendix C, we see the link between factoring a spectral density function and solving for C(L) from D(L) and H. Appealing to linear prediction theory and using the development provided in Appendix B it is possible to show that a whole family of H's and D(L)'s lead to the same decision rule. This turns out to be a simple corollary to the result that a covariance stationary stochastic process has multiple moving average representations.[13/] Thus, without further restrictions there is a whole family of objective functions that are consistent with decision rule (6.1). In absence of additional restrictions we cannot hope completely to identify the objective function parameters.

Fortunately, for many purposes, the fact that only a class of objective functions can be identified is of no practical concern. The reason is that all objective functions that imply the same decision rule give rise to exactly the same predictions about the response of economic agents to interventions in the form of changes in $\delta(L)$. For econometric policy evaluation, then, it is enough to identify the decision rule without having completely to identify the objective function.

In circumstances in which either more data are available or in which *a priori* restrictions are imposed on $D(L)$ it is often possible substantially to reduce the family of objective functions consistent with $C(L)$. For example, the econometrician may have observations on "output" q_t which obeys

$$(6.2) \qquad q_t = (h + S_{1t})'y_t - y_t'Hy_t$$

where S_1 is not observed by the econometrician. The idea here is that observations on q permit estimation of H via (6.2). In addition it is supposed that the cost term $[D(L)y_t]' \; [D(L)y_t]$ in 2.1 represents costs that are "internal" to the firm or the unit whose decisions are being modeled, so that $D(L)$ cannot be estimated from direct observations on inputs and outputs. In this example, H is uniquely identified; however, $D(L)$ cannot necessarily be completely pinned down from estimates of $C(L)$. A different example is where the form of $D(L)$ is restricted so that

$$[y_t - y_{t-1}]' \; \bar{H} \; [y_t - y_{t-1}] = [D(L)y_t]' \; [D(L)y_t] \quad .$$

Even without observations on q_t, it is possible to recover both $\bar{H}$ and H from $C(L)$ and hence the objective function parameters are all identified.

Conclusions

This paper has been devoted to describing quick and revealing ways of calculating optimal decision rules or dynamic equilibria for linear stochastic rational expectations models. The full value of such methods becomes evident only when we recall that our purpose is ultimately to estimate models of this class by using interpretations of the errors and estimators along the lines described in our earlier paper (Hansen and Sargent [7]). For example, there we describe maximum likelihood procedures for the estimation of single-endogenous variable, dynamic models of the class considered here. For the purposes of implementing maximum likelihood methods, it is a substantial advantage to have quick algorithms for evaluating the likelihood function, which requires evaluating the optimal decision rule or equilibrium stochastic process. It is also an advantage to have formulas as close to being in closed form as possible, since this facilitates computing analytic derivatives of the likelihood function. The general principles of estimation and interpretation of error terms described in our earlier paper extend in a fairly straightforward way to the present context.

Appendix A

In this appendix we examine the solutions to discrete time Euler equations for the infinite time problem. We take the following steps in order to characterize these solutions. First, write the Euler equations

$$\text{(A.1)} \qquad [H+D(\beta L^{-1})'D(L)]y_t = S_t^* \text{ for } t = 0, 1, \ldots.$$

Second, obtain a partial fractions decomposition of $[H+D(\beta z^{-1})'D(z)]^{-1}$ of the form

$$[H+D(\beta z^{-1})'D(z)]^{-1} = \frac{G_1^*}{z-z_1} + \ldots + \frac{G_k^*}{z-z_k} + \frac{H_1^*}{z-\beta z_1^{-1}} + \ldots + \frac{H_k^*}{z-\beta z_k^{-1}}$$

where $z_1, \ldots, z_k$ are assumed to be distinct and are greater than $\sqrt{\beta}$ in modulus. We assume that $(\frac{1}{\sqrt{\beta}})^t S_t^* \to 0$ as $t \to \infty$. Third, we obtain a particular solution to (A.1) of the form

$$y_t^p = \frac{-G_1^*}{z_1} \sum_{j=0}^{\infty} (z_1)^{-j} S_{t-j}^* - \ldots \frac{-G_k^*}{z_k} \sum_{j=0}^{\infty} (z_k)^{-j} S_{t-j}^* + H_1^* \sum_{j=0}^{\infty} (\beta z_1^{-1})^j S_{t+j+1}^* + \ldots + H_k^* \sum_{j=0}^{\infty} (\beta z_k^{-1})^j S_{t+j+1}^*$$

for $t = -m, -m+1, \ldots$ where $S_j^* = 0$ for $j < 0$. Let $A_1^*, \ldots, A_k^*, B_1^*, \ldots, B_k^*$ be n dimensional nonzero vectors such that

$$[H+D(\beta z_j^{-1})'D(z_j)]A_j^* = 0$$

$$[H+D(z_j)'D(\beta z_j^{-1})]B_j^* = 0.$$

The general solution to the homogenous equation is

$$y_t^h = c_1 A_1^* z_1^{-t} + \ldots + c_k A_k^* z_k^{-t} + f_1 B_1^* (\beta z_1^{-1})^{-t} + \ldots + f_k B_k^* (\beta z_k^{-1})^{-t}$$

where $c_1, \ldots, c_k, f_1, \ldots, f_k$ are arbitrary scaler constants. Fifth, obtain a general representation of the solutions to (A.1) by adding y_t^p and y_t^h.

We have m x n = k initial conditions $y_{-1}, y_{-2}, \ldots, y_{-m}$. We also have the requirement that

$$\text{(A.2)} \qquad \sum_{t=0}^{\infty} \beta^t y_t' H y_t < \infty$$

where H is positive definite. Note that for nonzero f_j

$$\sum_{t=0}^{\infty} \beta^t f_j^2 (\beta z_1^{-1})^{-2t} B_j^{*\prime} H B_j^*$$

is not finite. Thus, (A.2) is satisfied only if $f_j = 0$ for $j = 1, \ldots, k$. The initial condition vectors $y_{-1}, \ldots, y_{-m}$ uniquely determine $c_1, \ldots, c_k$. The solution provided in the text corresponds to the solution for y_t described above.

Appendix B

This appendix proves the assertions in the text about the factorization of the characteristic polynomial associated with the system of Euler equations. We state the assertions in the form of the following

Lemma: The matrix polynomial in z, $[H + D(\beta z^{-1})'D(z)]$, has a representation

$$H + D(\beta z^{-1})'D(z) = C(\beta z^{-1})'C(z) \tag{B1}$$

where $C(z) = \sum_{j=0}^{m} C_j z^j$, each C_j is an (nxn) matrix, and all the roots of det $C(z) = 0$ in modulus are not less than $\sqrt{\beta}$. The factorization (B1) is unique up to premultiplication of C(z) by an orthogonal matrix.

Proof: Define the polynomial in z

$$\begin{aligned} J(z) &= H + D(\beta z^{-1})'D(z) \\ &\equiv H + \bar{D}(\sqrt{\beta}\, z^{-1})'\, \bar{D}(\tfrac{1}{\sqrt{\beta}} z) \end{aligned} \tag{B2}$$

where $D(z) = \sum_{j=0}^{m} D_j z^j = \sum_{j=0}^{m} D_j (\sqrt{\beta})^j (\frac{1}{\sqrt{\beta}} z)^j$.

So we have defined $\bar{D}(z) = \sum_{j=0}^{m} \bar{D}_j z^j$ where $\bar{D}_j \equiv D_j(\sqrt{\beta})^j$. Notice that

$$\begin{aligned} D(\beta z^{-1})' &= \sum_{j=0}^{m} D_j'(\sqrt{\beta})^j (\sqrt{\beta}\; z)^{-j} \\ &= \bar{D}(\sqrt{\beta}\, z^{-1})' \;. \end{aligned}$$

Also $$D(z) = \bar{D}(\tfrac{1}{\sqrt{\beta}} z).$$

Now consider the function F(z) defined as

$$F(z) = H + \bar{D}(z^{-1})'\, \bar{D}(z) \;.$$

The function F(z) is the matrix cross covariance generating function of the n-dimensional covariance stationary stochastic process W defined by

$$W_t = Y_t + X_t$$

where

$$E\, Y_t Y'_{t-s} = \begin{cases} H & s=0 \\ 0 & s\neq 0 \end{cases}$$

$$E\, Y_t = 0$$

$$X_t = \bar{D}(L)' U_t$$

$$EU_t = 0$$

$$EU_t U'_t = I$$

$$E\, Y_t U_{t-s} = 0 \text{ for all } s.$$

It follows from the factorization theorem for spectral density matrices (see Roxanov [21]) that we have the factorization of F(z),

$$\text{(B3)} \qquad H + \bar{D}(z^{-1})' \bar{D}(z) = G(z^{-1})' G(z)$$

where the roots of G(z) do not lie inside the unit circle. The factorization is unique up to premultiplication of G(z) by an orthogonal matrix. It immediately follows from (B2) that

$$H + D(\beta z^{-1})' D(z) = G(\sqrt{\beta}\, z^{-1})'\, G(\frac{1}{\sqrt{\beta}} z)$$

$$\text{(B4)} \qquad \equiv C(\beta z^{-1})' C(z)$$

where $C(z) = \sum_{j=0}^{m} C_j z^j$

and

$$C_j = \left(\frac{1}{\sqrt{\beta}}\right)^j G_j \quad .$$

From the spectral factorization theorem we know that

$$\det G(z) = \mu_0(1-\mu_1 z)\ldots(1-\mu_s\ z)$$

where $s \triangleq nm$, and where $|\mu_j| \leq 1$. Thus,

(B5)
$$\det C(z) = \det G\left(\frac{1}{\sqrt{\beta}} z\right) = \mu_0\left(1-\mu_1 \frac{1}{\sqrt{\beta}} z\right)\ldots\left(1-\mu_s \frac{1}{\sqrt{\beta}} z\right).$$

From (B5) we know that the roots of $\det C(z)$ are not less than $\sqrt{\beta}$ in modulus. It also follows that the roots of $\det C(\beta z^{-1})$ do not exceed $\sqrt{\beta}$ in modulus. This concludes the proof of the lemma.

Appendix C

This appendix provides explicit formulas for factoring $H + \tilde{D}(z^{-1})'\tilde{D}(z)$ where $n = 2$ and $m = 1$. Let $H + \tilde{D}(z^{-1})'\tilde{D}(z) = F(z)$, where we write

$$F(z) = \begin{bmatrix} f_{11}(z) & f_{12}(z) \\ f_{21}(z) & f_{22}(z) \end{bmatrix}$$

where we let

$$f_{11}(z) = \alpha_1 z^{-1} + \alpha_0 + \alpha_1 z$$

$$f_{22}(z) = \beta_1 z^{-1} + \beta_0 + \beta_1 z$$

$$f_{12}(z) = \gamma_{-1} z^{-1} + \gamma_0 + \gamma_1 z$$

$$f_{21}(z) = \gamma_1 z^{-1} + \gamma_0 + \gamma_{-1} z$$

The factorization procedure involves the following three steps:

Step 1: Set $f_{11}(z) = \rho_0(1-\rho_1 z)(1-\rho_1 z^{-1})$, $|\rho_1| < 1$, $\rho_0 > 0$.

This is accomplished by setting

$$\rho_1 = \frac{-\alpha_0 \pm \sqrt{\alpha_0^2 - 4\alpha_1^2}}{2\alpha_1} \qquad \text{subject to } |\rho_1| < 1$$

$$\rho_0 = -\frac{\alpha_1}{\rho_1}$$

Step 2: Form det $F(z)$ and find the factorization

$$\det F(z) = \kappa_0(1-\kappa_1 z)(1-\kappa_2 z)(1-\kappa_1 z^{-1})(1-\kappa_2 z^{-1})$$

where $\kappa_0 > 0$, $|\kappa_1| < 1$, $|\kappa_2| < 1$. This is accomplished as follows. Let

$$a_2 = \alpha_1\beta_1 - \gamma_{-1}\gamma_1$$

$$a_1 = \alpha_0\beta_1 + \alpha_1\beta_0 - \gamma_0(\gamma_1+\gamma_{-1})$$

$$a_0 = \alpha_0\beta_0 + 2\alpha_1\beta_1 - \gamma_{-1}^2 - \gamma_1^2 - \gamma_0^2.$$

If $a_2 \neq 0$ set

$$\hat{\kappa}_1 = \frac{\frac{-a_1}{a_2} + \sqrt{\frac{a_1^2}{a_2^2} - 4(\frac{a_0}{a_2}-2)}}{2}$$

$$\hat{\kappa}_2 = \frac{\frac{-a_1}{a_2} - \sqrt{\frac{a_1^2}{a_2^2} - 4(\frac{a_0}{a_2}-2)}}{2}$$

$$\kappa_1 = \frac{\hat{\kappa}_1 \pm \sqrt{\hat{\kappa}_1^2 - 4}}{2} \qquad \text{subject to } |\kappa_1| < 1$$

$$\kappa_2 = \frac{\hat{\kappa}_2 \pm \sqrt{\hat{\kappa}_2^2 - 4}}{2} \qquad \text{subject to } |\kappa_2| < 1$$

$$\kappa_0 = \frac{a_2}{\kappa_1\kappa_2}$$

If $a_2 = 0$ set

$$\kappa_1 = 0$$

$$\kappa_2 = \frac{-a_0 \pm \sqrt{a_0^2 - 4a_1^2}}{2a_1} \qquad \text{subject to } |\kappa_2| < 1$$

$$\kappa_0 = -\frac{a_2}{\kappa_2}.$$

<u>Step 3</u>: Compute $G(z)$ where $G(z^{-1})'G(z) = F(z)$, where

$$G(z) = \begin{bmatrix} g_{11}(z) & g_{12}(z) \\ g_{21}(z) & g_{22}(z) \end{bmatrix}$$

We compute G(z) as follows.

Define

$$\xi_0 = \gamma_{-1} + \gamma_0\rho_1 + \gamma_1\rho_1^2,$$

$$\xi_1 = \sqrt{\kappa_0}(1-\kappa_1\rho_1)(1-\kappa_2\rho_1)$$

$$\xi_2 = \sqrt{\xi_0^2 + \xi_1^2}$$

$$\psi_0 = \xi_0/\xi_2$$

$$\psi_1 = \xi_1/\xi_2$$

$$\kappa_3 = -(\kappa_1+\kappa_2)$$

Then

$$g_{11}(z) = \psi_0\sqrt{\rho_0}(z-\rho_1)$$

$$g_{21}(z) = -\psi_1\sqrt{\rho_0}(1-\rho_1 z)$$

$$g_{12}(z) = \frac{\psi_0\gamma_{-1} + \psi_1\sqrt{\kappa_0} + (\psi_0\gamma_0+\rho_1\psi_0\gamma_{-1}+\rho_1\psi_1\sqrt{\kappa_0}+\psi_1\sqrt{\kappa_0}\kappa_3)z}{\sqrt{\rho_0}}$$

$$g_{22}(z) = \frac{\psi_1\gamma_{-1}-\psi_0\sqrt{\kappa_0} + (\psi_1\gamma_0+\psi_1\gamma_{-1}\rho_1^{-1}-\psi_0\sqrt{\kappa_0}\kappa_3-\psi_0\sqrt{\kappa_0}\rho_1^{-1})z}{\sqrt{\rho_0}\rho_1}$$

Footnotes

1/ The model could be extended in various ways to model the determination of J_t and W_t in terms of the demands for factors generated here interacting with supply curves for these factors.

2/ To insure that there is a sense in which the criterion function is well defined as $N\to\infty$, we could impose the weak restriction on $\sum_t$ that

$$\lim_{t\to\infty} \sup \beta^t \text{ trace } \textstyle\sum_t = 0$$

It turns out that even this weak restriction on $\sum_t$ could be relaxed if we adopt a suitable modification of the criterion function (2.1) below. For example, replacing (2.1) with

$$\lim_{N\to\infty} \frac{1}{N} E_0 \sum_{t=0}^{N} \beta^t \{(h+S_{1t})' y_t - y_t' H y_t - [D(L) y_t]' [D(L) y_t]\},$$

would yield the same decision rule, but permit weakening the above condition on $\sum_t$.

3/ For problems with negative semidefinite matrixes R, the "appropriate" solution of (2.5) is the unique negative definite solution P. For our problem, R fails to be negative semidefinite, and so does the appropriate solution of (2.5). However, the sub-matrix R_{11} defined below is negative semidefinite, and so is the associated P_{11}. This is enough to make our problem well posed, and to support the claims about appropriate solutions that are made in the text. These claims are proved in Sargent [22].

4/ For details, see Sargent [22].

5/ This can be established as follows. In (2.3), define the transformed variables $\tilde{X}_t = \beta^{t/2} X_t$, and $\tilde{v}_t = \beta^{t/2} v_t$. Problem (2.3) is equivalent to the undiscounted linear regulator problem, to maximize

$$\lim_{N\to\infty} E_0 \sum_{t=0}^{N} \{\tilde{X}_t'R\tilde{X}_t+\tilde{v}_t'Q\tilde{v}_t\}$$

subject to $\tilde{X}_{t+1} = \beta^{1/2}AX_t + \beta^{1/2}B\tilde{v}_t + \beta^{(t+1)/2}\xi_{t+1}$. The optimizer is a control law $\tilde{v}_t = \tilde{F}\tilde{X}_t$, from which the optimizer of the original problem $v_t = \tilde{F}X_t$ can be calculated. For A, B, R, and Q defined as in the text, it can be verified that the pair $(\beta^{1/2}A,\beta^{1/2}B)$ is stabilizable. Further, letting the rank of R be r, choose a matrix G with r rows such that $G'G = R$. Then for our problem, it can be verified that the pair $(\beta^{1/2}A,G)$ is detectable. The stabilizability of $(\beta^{1/2}A, \beta^{1/2}B)$ and the detectability of $(\beta^{1/2}A,G)$ imply that the closed loop system matrix $(\beta^{1/2}A-\beta^{1/2}B\tilde{F})$ has all of its eigenvalues less than unity in modulus. From this and the observation that $\tilde{F} = F$, it follows that the eigenvalues of (A-BF) are bounded by $\beta^{-1/2}$ in modulus. For a more detailed treatment, see Sargent [22]. Kwakernaak and Sivan [12] and Kailath [9] are good references on the results from linear optimal control theory we are appealing to.

6/ Necessary and sufficient conditions for iterations on (2.6) to converge are readily stated in terms of the controllability canonical form (2.9). The necessary and sufficient conditons are (a), that the pair (A_{11},B_1) be controllable, and (b) that the eigenvalues of the matrix A_{22} have moduli all less than $1/\sqrt{\beta}$. The necessary and sufficient condition for controllability is that the matrix $[B_1,A_{11}B_1,\ldots,A_{11}^{mn-1}B_1]$ have rank mn, where m is the number of lags in D(L) and n is the dimension of y_t. Condition (b) on the eigenvalues of A_{22} is guaranteed by our assumption that the roots of det $\delta(z) = 0$ are all greater than $\sqrt{\beta}$ in modulus. These conditions are derived mainly by adapting results summarized by Kwakernaak and Sivan [9] for the undiscounted case to the discounted case. See Sargent [22] for details.

7/ Again, for details see Sargent [22].

8/ The assumption that H is positive definite, and not only semidefinite, plays a role in delivering the bound of $\beta^{-1/2}$ on the modulus of the eigenvalues of $A_{11}-B_1F_1$. If H were only assumed to be positive semidefinite, more restrictions than have been imposed above on D(L) would be required to assure that the eigenvalues of $(A_{11}-B_1F_1)$ are bounded in modulus by $\beta^{-1/2}$. In particular, sufficient conditions would have to be imposed to satisfy the detectability condition described in footnote 5.

9/ This follows from the stabilizability and detectability of the transformed system (see footnote 5) and from results in linear optimal control theory (see Kwakernaak and Sivan [12] and Sargent [22]).

10/ The procedure suggested by Rozanov [21] involves obtaining an intial non-invertible triangular factorization and then multiplying by appropriate Blaschke factor matrices in order to shift roots from inside the unit circle to outside the unit circle. An alternative procedure for factoring a vector moving average, discrete time spectral density matrix, has been offered by Whittle [28] and more recently by Murthy [19]. It amounts to inverting the spectral density, thus converting a vector moving average problem into a vector autoregressive problem. The orthogonality conditions associated with the vector autoregression are then used to determine the invertible factorization. From the standpoint of this paper, we are concerned only with the factorization of spectral densities of vector moving average processes. However, the procedures discussed by Rozanov, Whittle, and Murthy are appropriate for arbitrary rational spectral densities.

11/ Note that in this example the matrix $H = A_1 m^2 dd'$, and so is positive semidefinite but not positive definite. However, it can be verified that the problem does satisfy sufficient conditions for the closed loop system matrix $(A_{11}-B_1F_1)$ to have eigenvalues bounded in modulus by $\beta^{-1/2}$. In particular, the zeroes of what corresponds to the matrix polynomial det D(z) of Section 3 are less than $\beta^{-1/2}$ in modulus, which delivers the required detectability condition (see footnote 5).

12/ We did not carry along the constant terms in the demand function or compute them for the equilibrium. The latter constants would be easy to compute given the former.

13/ Multiple moving average representations can be obtained both by "flipping" roots inside and outside the unit circle via multiplication by Blaschke factors and by altering the number of underlying orthogonal white noise processes employed in the representation. See Rozanov [21] for details.

References

[1] Bertsekas, Dimitri P., Dynamic Programming and Stochastic Control, Academic Press, New York, 1976.

[2] Emre, Erol, and Özay Hüseyin, "Generalization of Leverrier's Algorithm to Polynomial Matrices of Arbitrary Degree," IEEE Transaction Automatic Control, February 1975, p. 136.

[3] Epple, D., and L. Hansen, "An Econometric Model of U.S. Petroleum Supply with Optimal Endogenous Depletion," unpublished manuscript, April 1979.

[4] Gould, J. P., "Adjustment Costs in the Theory of Investment of the Firm," The Review of Economic Studies, 1968, Vol. XXXV(1), No. 101, pp. 47-56.

[5] Granger, C. W. J., "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," Econometrica 37 (July 1969): pp. 424-438.

[6] Graves, R., and L. Telser, Functional Analysis in Mathematical Economics, Chicago: University of Chicago Press, 1971.

[7] Hansen, Lars P., and Thomas Sargent, "Formulating and Estimating Dynamic Linear Rational Expectations Models," Journal of Economic Dynamics and Control (forthcoming), 1980.

[8] Holt, C., F. Modigliani, T. F. Muth, and H. A. Simon, Planning Production, Inventories and Work Force, Englewood Cliffs, New Jersey: Prentice Hall, 1960.

[9] Kailath, Thomas, Linear Systems, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1979.

[10] Kennan, John, "The Estimation of Partial Adjustment Models with Rational Expectations," Econometrica (forthcoming), 1980.

[11] Kushner, Harold, Introduction to Stochastic Control, New York: Holt Rinehart and Winston, 1971.

[12] Kwakernaak, H., and R. Sivan, Linear Optimal Control Systems, New York: Wiley, 1972.

[13] Lucas, R. E., Jr., "Optimal Investment Policy and the Flexible Accelerator," International Economic Review, Vol. 8, No. 1, February 1967; pp. 78-85.

[14] ____________, "Adjustment Costs and the Theory of Supply," Journal of Political Economy, 1967, Part 1, Vol. 75, No. 4, pp. 321-334.

[15] ____________, "Econometric Testing of the Natural Rate Hypothesis," in <u>The Econometrics of Price Determination Converence</u>, edited by Otto Eckstein, Washington: Board of Governors, Federal Reserve System, 1972.

[16] ____________, "Econometric Policy Evaluation: A Critique," in <u>The Phillips Curve and Labor Markets</u>. Edited by K. Brunner and A. H. Meltzer, Carnegie-Rochester Conference Series on Public Policy 1, Amsterdam: North Holland, 1976, pp. 19-46.

[17] ____________, and E. Prescott, "Investment Under Uncertainty," <u>Econometrica</u> 39, September 1971: pp. 659-681.

[18] Mortensen, Dale, "Generalized Costs of Adjustment and Dynamic Factor Demand Theory," <u>Econometrica</u> 41, 1973, pp. 657-665.

[19] Murthy, D. N. P., "Factorization of Discrete-Process Spectral Matrices," <u>IEEE Transactions on Information Theory</u>, September 1973, pp. 693-696.

[20] Nadiri, M., and S. Rosen, <u>A Disequilibrium Model of Demand for Factors of Production</u>, New York: Columbia University Press, 1973.

[21] Rozanov, Yu. A., <u>Stationary Random Processes</u>, Holden-Day, San Francisco, 1967.

[22] Sargent, Thomas J., "Lecture Notes on Linear Prediction and Control," unpublished manuscript, Winter 1980.

[23] ____________, <u>Macroeconomic Theory</u>, Academic Press, 1979, Chapter XIV.

[24] Taylor, John B., "Output and Price Stability: An International Comparison," manuscript June 1978.

[25] ____________, "Estimation and Control of a Macroeconomic Model with Rational Expectations," <u>Econometrica</u> (forthcoming).

[26] Treadway, Arthur B., "On Rational Entrepreneurial Behavior and the Demand for Investment," <u>Review of Economic Studies</u>, 1969, Vol. XXXVI(2), No. 106, pp. 227-240.

[27] ____________, "The Rational Multivariate Flexible Accelerator," <u>Econometrica</u>, Vol. 39, No. 5, September 1971, pp. 845-855.

[28] Whittle, P., <u>Regulation and Prediction by Linear Least Squares Methods</u>, Princeton: D. Van Nostrand, 1973.