7

Faster Methods for Solving Continuous Time Recursive Linear Models of Dynamic Economies

by Lars Peter HANSEN, John HEATON and Thomas J. SARGENT

Introduction

This paper describes calculations that are useful for computing equilibria of recursive linear models of general competitive equilibrium in continuous time. The calculations are designed to make it possible to implement continuous time reformulations of a class of models whose discrete time versions are analyzed by Hansen and Sargent (1990). As in Hansen and Sargent's discrete time setting, the basic idea is to formulate a fictitious social planning problem whose solution equals the allocation that would be associated with the competitive equilibrium of a decentralized version of the economy. The social planning problem has a quadratic objective function and linear constraints. Some of the constraints are represented as non-autonomous differential equations, which means that we permit the endowment and preference shocks to be relatively general functions of time. In particular, the preference and endowment shocks need not be restricted to be themselves the outputs of a system of differentiable equations (this is what we mean by saying that the constraints may take the form of *nonautonomous* differential equations). Permitting the preference and endowment shocks to have this general structure is designed eventually to facilitate a number of applications that we have in mind. The intention is to formulate the social planning problem with ample generality by allowing room to include a variety of setups with potentially large numbers of capital stocks and informational state variables. As in Hansen and Sargent (1990), forcing the planning problem into the form of a linear-quadratic optimization problem yields substantial computational benefits: the equilibrium allocations can be computed by solving that optimization problem (which is often called an *optimal linear regulator problem*), and the Arrow-Debreu prices that support those allocations as competitive equilibria can be computed from information contained in the value function for that optimization problem.

There are standard methods available for solving a class of problems similar to ours known as optimal linear regulator problems (see Kwakernaak and Sivan 1972 and Anderson 1978). However, these algorithms require that we can represent the constraints as systems of autonomous differential equations. This paper shows how those standard methods can be adapted to handle problems in which some of the constraints take the form of nonautonomous differential equations. We proceed recursively, first solving a reduced system whose constraints are represented as an autonomous system of differential equations. In solving this reduced problem, we use fast algorithms to compute the *feedback* part of the solution of the original problem. Second, we compute an additional part of the solution, called the *feedforward* part, which incorporates the effect of the nonautonomous parts of the differential equations describing the constraints. Fast algorithms for implementing both parts of these calculations are described in this paper.

This paper can be viewed as an extended technical prolegomenon to future work that will describe a continuous-time version of Hansen-Sargent (1990). In effect, we describe continuous time versions of an array of computational tricks that have discrete time analogues that are studied by Hansen and Sargent (1990). We concentrate our attention on solving the social planning problem, and do not lay out the connection of the social planning problem to a continuous time version of a competitive equilibrium. However, it is important to note that the connection to a competitive equilibrium is there, and that it can be spelled out by mimicking in continuous time versions of the arguments that Hansen and Sargent utilize in discrete time.

This paper studies a nonstochastic version of the model. However, just as in Hansen-Sargent (1990), once the solution of the planning problem under certainty is obtained, it is a simple matter to compute the equilibrium of corresponding stochastic economies. This is a standard feature of economies whose planning problems are optimal linear regulators. In the next paper in this volume, a set of continuous time prediction formulas that have been devised to compute the *feedforward* parts of a stochastic version of our models. As emphasized by Lucas and Sargent (1981) a convenient feature of linear quadratic models is the manner in which their solution *separates* into a control part and a predition part. This paper and the next one correspond to this separation.

This paper is organized as follows. Section 1 sets forth the social planning or optimal resource allocation problem. The planning problem assumes the form of an optimal linear regulation problem. Section 2 defines several linear operators that are useful in deriving and representing a solution. Section 3 uses some of these operators to represent the constraints impinging on the planning problem in a way that facilitates computing a solution. Section 4 solves our particular optimal linear regulator problem in a way that exploits its special structure. Section 5 describes how a *matrix sign* algorithm can be applied to simplify and speed aspects of the computations. Sections 6 and 7 are devoted to studying two special cases of our model. A class of continuous time models of costly adjustment of capital are described in Section 6, while Section 7 studies a version of Heaton's (1989) model of consumer durables.

1. Optimal Resource Allocation Problem

We are interested in computing solutions to an optimal resource allocation problem that is the continuous-time counterpart to a discrete-time problem studied extensively by Hansen and Sargent (1990). The solution to this optimization problem can be decentralized and interpreted as the time path for the equilibrium quantities of an intertemporal competitive equilibrium.

Consider the following general setup. A household has time separable preferences defined over an n_s -dimensional vector $s(t)$ of *services* at time t. The preferences are given by:

$$(1.1) \qquad -(1/2)\int_0^\infty \exp(-\rho t)\ [s(t)-b(t)]'\ [s(t)-b(t)]dt$$

where $b(t)$ is an exogenous n_s -dimensional vector function of time and $\rho > 0$ is the subjective rate of discount. Services at time t are generated according to:

$$(1.2) \qquad s(t) = \Lambda h(t) + \Pi c(t)$$

where $h(t)$ is a n_h -dimensional vector of *household capital* stocks at time t, $c(t)$ is an n_c -dimensional vector of *consumption goods* at time t, Λ is an $n_s \times n_h$ matrix and Π is an $n_s \times n_c$ matrix.

The household capital stock evolves according to the following system of linear differential equations:

$$(1.3) \qquad Dh(t) = \Delta_h\, h(t) + \Theta_h\, c(t), \quad \text{for} \quad t \geq 0,\ h(0) = \mu_h\ ,$$

where Δ_h is an $n_h \times n_h$ matrix, Θ_h is an $n_h \times n_c$ matrix and μ_h is an $n_h \times 1$ vector of initial conditions. The matrix Δ_h governs the depreciation of the household capital stock.

Taken together, (1.2) and (1.3) induce intertemporal nonseparabilities into the indirect preference ordering for consumption. Among other things, the matrices Λ and Π determine the extent to which the consumption goods are substitutes or complements over time.

To allow the transfer of output over time, there is an n_k vector $k(t)$ consisting of *physical capital* stocks at time t. Capital is assumed to evolve over time according to:

$$(1.4) \qquad Dk(t) = \Delta_k\, k(t) + \Theta_k\, i(t), \quad \text{for} \quad t \geq 0, \; k(0) = \mu_k$$

where $i(t)$ is an n_i vector of *investment goods* at time t, Δ_k is an $n_k \times n_k$ matrix, Θ_k is an $n_k \times n_i$ matrix and μ_k is an $n_k \times 1$ vector of initial conditions. The matrix Δ_k determines the depreciation of the physical capital stocks.

Output is produced using the current capital stock and an endowment $f(t)$. This output is then divided between consumption and investment goods according to:

$$(1.5) \qquad \Phi_c\, c(t) + \Phi_i\, i(t) = \Gamma\, k(t) + f(t)$$

where Φ_c is an n_c -dimensional nonsingular matrix, Φ_i is an $n_c \times n_i$ matrix and Γ is an $n_c \times n_k$ matrix.

The functions $f(t)$ and $b(t)$ are driven by an n_z -dimensional vector of forcing functions of time denoted $\hat{z}(t)$:

$$(1.6) \qquad f(t) = \Xi_f\, \hat{z}(t) \quad \text{and} \quad b(t) = \Xi_b\, \hat{z}(t)$$

where Ξ_f is an n_c by n_z matrix and Ξ_b is an n_s by n_z matrix.

Suppose now that we define a composite state vector $\hat{x}(t)' \equiv [h(t)'\ k(t)']$ and a control vector $\hat{u}(t) \equiv i(t)$. Using the resource constraint (1.5) and the differential equation systems (1.3) and (1.4), we get the following system of differential equations governing $\hat{x}(t)$:

$$(1.7) \qquad D\hat{x}(t) = \hat{A}\hat{x}(t) + B_u\, \hat{u}(t) + B_z \hat{z}(t) \qquad \text{subject to} \;\; \hat{x}(0) = \mu$$

where $\hat{A} \equiv \begin{bmatrix} \Delta_h & \Theta_h(\Phi_c)^{-1}\Gamma \\ 0 & \Delta_k \end{bmatrix}$, $B_u \equiv \begin{bmatrix} -\Theta_h(\Phi_c)^{-1}\Phi_i \\ \Theta_k \end{bmatrix}$, $B_z \equiv \begin{bmatrix} \Theta_h(\Phi_c)^{-1}\Xi_f \\ 0 \end{bmatrix}$ and $\mu' \equiv (\mu_h', \mu_k')$. Notice that

$$(1.8) \qquad s(t) - b(t) = \Sigma_u\, \hat{u}(t) + \Sigma_x \hat{x}(t) + \Sigma_z \hat{z}(t)$$

where $\Sigma_u \equiv -\Pi(\Phi_c)^{-1}\Phi_i$, $\Sigma_x \equiv [\Lambda\ ;\ \Pi(\Phi_c)^{-1}\Gamma]$ and $\Sigma_z \equiv \Pi(\Phi_c)^{-1}\Xi_f - \Xi_b$. We can now write the objective function (1.1) in terms of $\hat{u}(t)$, $\hat{x}(t)$ and $\hat{z}(t)$:

$$(1.9) \qquad -(1/2)\int_0^\infty \exp(-\rho t)\; [\hat{u}(t)'\ \hat{x}(t)'\ \hat{z}(t)']\, \Omega \begin{bmatrix} \hat{u}(t) \\ \hat{x}(t) \\ \hat{z}(t) \end{bmatrix} dt$$

where $\Omega \equiv \begin{bmatrix} \Omega_{uu} & \Omega_{ux} & \Omega_{uz} \\ \Omega_{ux}' & \Omega_{xx} & \Omega_{xz} \\ \Omega_{uz}' & \Omega_{xz}' & \Omega_{zz} \end{bmatrix}$. The partitions of Ω are given by $\Omega_{uu} = \Sigma_u'\Sigma_u$, $\Omega_{ux} = \Sigma_u'\Sigma_x$, and so on. Notice that Ω is positive semidefinite.

We assume that the capital stocks and forcing functions satisfy:

$$(1.10) \qquad \begin{aligned} &\int_0^\infty \exp(-\rho t) \mid h(t) \mid^2 dt < \infty, \quad \int_0^\infty \exp(-\rho t) \mid k(t) \mid^2 dt < \infty \\ &\int_0^\infty \exp(-\rho t) \mid \hat{z}(t) \mid^2 dt < \infty\, . \end{aligned}$$

These constraints limit growth in the respective functions of time. In the case of the capital stocks, these constraints are used instead of nonnegativity constraints because they are easier to impose. The plausibility of the resulting solutions can be checked in practice by solving the problem numerically subject to these constraints and checking that the resulting time paths for the capital stocks are nonnegative.

Prior to solving the model, it is convenient to transform the optimization problem to remove the discounting. This is done as follows. Let $\epsilon = \rho/2$ and define

$$(1.11) \qquad \begin{aligned} x(t) &\equiv \exp(-\epsilon t)\hat{x}(t), \quad z(t) \equiv \exp(-\epsilon t)\hat{z}(t) \\ \text{and} \quad u(t) &\equiv \exp(-\epsilon t)\hat{u}(t)\, . \end{aligned}$$

In terms of the transformed vector $x(t)$, the restrictions in (1.10) imply

$$(1.10') \qquad \int_0^\infty \mid x(t) \mid^2 dt < +\infty.$$

Notice that

$$(1.12) \qquad Dx(t) = -\epsilon x(t) + \exp(-\epsilon t) D\hat{x}(t)\, .$$

As a consequence $x(t)$ satisfies the following system of differential equations:

$$(1.13) \qquad Dx(t) = Ax(t) + B_u u(t) + B_z z(t)\, , \quad \text{for} \;\; t \geq 0, \; x(0) = \mu$$

where $A = \hat{A} - \epsilon I$. Also the objective function (1.9) can be written as:

$$(1.14) \qquad -(1/2)\int_0^\infty [u(t)'\ x(t)'\ z(t)']\ \Omega \begin{bmatrix} u(t) \\ x(t) \\ z(t) \end{bmatrix} dt\ .$$

Our problem now is to maximize (1.14) subject to (1.10′) and (1.13) by choosing $\{u(t)\}_{t=0}^{\infty}$.

2. Important Linear Operators

Let C denote the space of complex numbers and R the space of real numbers. Also, let L_1 denote the collection of Borel measurable functions x mapping R into the space of complex numbers C such that

$$(2.1) \qquad \int_R |\ x(t)\ |\ dt < \infty$$

and let L_2 denote the collection of Borel measurable functions x such that

$$(2.2) \qquad \int_R |\ x(t)\ |^2\ dt < \infty\ .$$

Let L_1^n and L_2^n be the spaces of all n-dimensional vectors of functions with components in L_1 and L_2 respectively. In light of (1.10′), the vector of functions x given in (1.11) can be restricted to be elements of L_2^n where $n = n_h + n_k$.

It is convenient to analyze four operators that map L_2^n into itself using Fourier methods. Among other things, these operators will be useful in characterizing a particular solution to the differential equation system (1.13). Our strategy for constructing these operators is as follows. First, we display the operators evaluated at functions in $L_1^n \cap L_2^n$. Second, we verify that the operators map $L_1^n \cap L_2^n$ into L_2^n. Consequently, we can compute Fourier transforms of the resulting functions. The Fourier transforms of the operators, i.e., the composition of the Fourier transform operator with the original operators, have simple extensions to all of L_2^n. The original operators can then be extended to all of L_2^n by taking inverse Fourier transforms. These steps will be made clear in our examples.

The first operator is an *aggregation over time* operator. It aggregates x over γ time units. It is defined on $L_1^n \cap L_2^n$ via

$$(2.3) \qquad U_\gamma(x)\,(t) = \int_o^\gamma x(t+\tau)d\tau$$

where x is in $L_1^n \cap L_2^n$. It is straightforward to show that $U_\gamma(x)$ is in L_1^n. For any function x in L_1^n we define the Fourier transform operator T evaluated at x to be:

$$(2.4) \qquad T(x)\,(\theta) = \int_R \exp(-i\theta t)\,x(t)dt$$

We evaluate the Fourier transform of $U_\gamma(x)$ by changing orders of integration. This gives:

$$(2.5) \qquad \begin{aligned} T[U_\gamma(x)](\theta) &= \int_R \exp(-i\theta t)\ [\int_0^\gamma x(t+\tau)d\tau]dt \\ &= \int_0^\gamma \exp(i\theta\tau)\ [\int_R \exp[-i\theta(t+\tau)]x(t+\tau)dt]d\tau \\ &= \{[\exp(i\theta\gamma) - 1]/(i\theta)\}T(x)(\theta)\ . \end{aligned}$$

The function ϕ, where $\phi(\theta) \equiv [\exp(i\theta\gamma) - 1]/(i\theta)$, is continuous in θ on R including zero. In addition, $\phi(\theta)$ goes to zero as $|\ \theta\ |$ goes to infinity. Therefore, ϕ is bounded in θ. It follows from a multidimensional analog to the Parseval formula that $U_\gamma(x)$ is in $L_1^n \cap L_2^n$ (e.g. see Rudin 1974, Chapter 9).

Notice that the operator obtained by multiplying x in L_2^n by ϕ maps L_2^n into L_2^n. We extend the construction of U_γ to all of L_2^n by letting $U_\gamma(x)$ be the element of L_2^n that has $\phi T(x)$ as its Fourier transform. In Appendix A we establish that this construction of U_γ is consistent with (2.3) and (2.5) for all x in L_2^n.

The second operator we consider is the *shift* operator. For any x in $L_1^n \cap L_2^n$ we define

$$(2.6) \qquad S_\gamma(x)\,(t) = x(t+\gamma)\ .$$

Taking Fourier transforms of both sides of (2.6) gives

$$(2.7) \qquad \begin{aligned} T[S_\gamma(x)](\theta) &= \int_R \exp(-i\theta t)x(t+\gamma)dt \\ &= \exp(i\theta\gamma)\int_R \exp[-i\theta(t+\gamma)]x(t+\gamma)dt \\ &= \exp(i\theta\gamma)T(x)(\theta)\ . \end{aligned}$$

These calculations can be extended from $L_1^n \cap L_2^n$ to L_2^n just as in the previous case except that now $\phi(\theta) \equiv \exp(i\theta\gamma)$. Note that $|\ \phi(\theta)\ | = 1$ for all θ. In Appendix A we show that this extension is compatible with (2.6).

The third operator is a *matrix backward convolution* operator. Again let x be in $L_1^n \cap L_2^n$, and let A be an $n \times n$ matrix of real numbers with eigenvalues that have strictly negative real parts. The *backward* convolution operator is then given by:

(2.8) $$C(x)(t) = \int_0^\infty \exp(A\tau)x(t-\tau)d\tau$$

where

(2.9) $$\exp(A) = \Sigma_{j=0}^{\infty} A^j/j\ !$$

The components of $C(x)$ are in L_1 since the eigenvalues of A have strictly negative real parts and

(2.10) $$\begin{aligned}\int_R |C(x)(t)|dt &\le \int_R \int_0^\infty |\exp(A\tau)|\ |x(t-\tau)|\, d\tau dt \\ &= \int_0^\infty |\exp(At)|\, dt\ \int_R |x(t)|\, dt\ .\end{aligned}$$

Hence formula (2.4) implies that

(2.11)
$$\begin{aligned}T[C(x)](\theta) &= \int_R \exp(-i\theta t)[\int_0^\infty \exp(A\tau)x(t-\tau)d\tau]dt \\ &= \int_0^\infty \exp(-i\theta\tau)\exp(A\tau)\,\{\int_R \exp[-i\theta(t-\tau)x(t-\tau)dt\}d\tau \\ &= (i\theta I - A)^{-1}T(x)(\theta)\ .\end{aligned}$$

Notice that $i\theta I - A$ is nonsingular for all θ in R because the eigenvalues of A have strictly negative real parts. Consequently,

(2.12) $$\bar{\delta} \equiv \sup_{\theta \in R} \text{trace}\,\{(-i\theta I - A')^{-1}(i\theta I - A)^{-1}\} < \infty\ .$$

Applying the Cauchy-Schwarz Inequality gives

(2.13) $$\int_0^\infty |\ T[C(x)(\theta)]\ |^2\, d\theta \le \bar{\delta} \int_0^\infty |\ T(x)(\theta)\ |^2\, d\theta\ .$$

Thus $T[C(x)]$ has components in L_2^n, and we extend C to all of L_2^n so that (2.11) is satisfied. In Appendix A we show that this extension is compatible with (2.8).

Using the same logic, we can define a *forward convolution* operator C' such that for any x in L_2^n,

(2.14) $$T[C'(x)](\theta) = (-i\theta I - A')^{-1}T(x)(\theta)\ .$$

When x is also in L_1^n, the transform operator has a time domain representation:

(2.15) $$C'(x)(t) = \int_0^\infty \exp(A'\tau)x(t+\tau)d\tau\ .$$

Thus the transform operator is a *forward* convolution operator. The operator $C'(x)$ can also be extended to all of L_2^n in a manner that preserves (2.15).

3. Linear Constraints

In Section 1, we described an optimization problem that is subject to a linear constraint that can be represented as a system of differential equations:

(3.1) $$Dx(t) = Ax(t) + Bv(t) \quad \text{for} \quad t \ge 0,\ x(0) = \mu$$

where $B = [B_u\ B_z]$ and $v(t)' = [u(t)'\ z(t)']$. We refer to equation (3.1) as the *state equation*. In this section we describe how to solve the state equation and obtain a frequency domain representation of the constraints. As in Section 2, we restrict the eigenvalues of A:

Assumption 3.1: The eigenvalues of the matrix A have real parts that are strictly negative.

If the matrices Δ_h and Δ_k discussed in Section 1 all have eigenvalues with real parts that are strictly less than ϵ, then the model of Section 1 will satisfy Assumption 3.1.

We restrict v to be in L_2^N where $N = n_i + n_z$. For convenience, we set $v(t)$ to zero for strictly negative values of t. The space of all elements of L_2^N that are zero for negative t will be denoted L_+^N. We now consider solving (3.1) for x in terms of v in L_+^N. We deduce a particular solution to (3.1) that applies for all time periods but the solution will not necessarily satisfy $x(0) = \mu$. The particular solution will be denoted x^p.

For equation (3.1) to be satisfied for all t, x^p must satisfy:

(3.2) $$x^p(t+\gamma) - x^p(t) = A\int_t^{t+\gamma} x^p(s)ds + B\int_t^{t+\gamma} v(s)ds$$

for all t and all strictly positive values of γ. Expressing (3.2) in operator notation gives

(3.3) $$S_\gamma(x^p) - x^p = AU_\gamma(x^p) + BU_\gamma(v)\ .$$

Taking Fourier transforms gives

$$(3.4)\quad \begin{aligned}&[\exp(i\theta\tau)-1]T(x^p)(\theta)\\ &\quad = \{[\exp(i\theta\tau)-1]/(i\theta)\}\,[AT(x^p)(\theta)+BT(v)(\theta)]\,.\end{aligned}$$

Solving (3.4) for $T(x^p)$ gives

$$(3.5)\qquad T(x^p)(\theta) = (i\theta I - A)^{-1}BT(v)(\theta)\,.$$

From the analysis in Section 2, we know that $x^p = C(Bv)$, or

$$(3.6)\qquad x^p(t) = \int_0^\infty \exp(A\tau)Bv(t-\tau)d\tau\,,$$

[see (2.11)]. Equations (3.1) and (3.5) suggest a convenient notation for the convolution operator that is analogous to the notation used for lag operators in discrete time. The backward convolution operator applied to an n-dimensional vector of functions, x, will be written as:

$$(3.7)\qquad C(x)(t) \equiv (DI - A)^{-1}x(t)\,.$$

where we have replaced $i\theta$ by D.

Using this new notation, we represent x^p as:

$$(3.8)\qquad x^p(t) = (DI - A)^{-1}Bv(t)\,.$$

Since $v(t)$ is zero for strictly negative values of t, it follows that x^p is in L_+^n and that for $t \geq 0$

$$(3.9)\qquad x^p(t) = \int_0^t \exp(A\tau)Bv(t-\tau)d\tau\,.$$

The process $\{x^p(t) : -\infty < t < +\infty\}$ given by (3.9) does not, in general, satisfy the initial condition because (3.9) implies that $x^p(0)$ is zero. The homogeneous differential equation

$$(3.10)\qquad Dx^h(t) = Ax^h(t)\ \text{ for }\ t \geq 0,\ x(0) = \mu$$

has as its solution

$$(3.11)\qquad x^h(t) = \exp(At)\mu\,.$$

This can be verified by differentiating the power series expansion for $\exp(At)$ with respect to t.

Adding together x^p to x^h gives a solution to the original differential equation (3.1). Hence

$$(3.12)\qquad x(t) = \exp(At)\mu + \int_0^t \exp(A\tau)Bv(t-\tau)d\tau\ \text{ for }\ t \geq 0\,.$$

Equality (3.12) will be taken as the constraint for the optimization problem we study in Section 4.

There is a differential equation system that is closely related to equation system (3.1). Let x and λ be processes in L_2^n such that

$$(3.13)\qquad Dx(t) = -A'x(t) - \lambda(t)\,.$$

Notice that the eigenvalues of $-A'$ have real parts that are strictly positive. The function λ will be a Lagrange multiplier in our analysis in Section 4.

We proceed as before by using Fourier transform methods to solve (3.13) for x in terms of λ. Expressed in terms of transforms, this gives

$$(3.14)\qquad T(x)(\theta) = (-i\theta I - A')^{-1}T(\lambda)(\theta)\,.$$

It follows from the analysis in Section 2 that:

$$(3.15)\qquad x(t) = C'(\lambda)(t) = \int_0^\infty \exp(A'\tau)\lambda(t+\tau)d\tau\,.$$

Just as in the case of the backward convolution operator, (3.13) and (3.15) suggest a convenient notation for the forward convolution operator:

$$(3.16)\qquad C'(\lambda)(t) \equiv (-DI - A')^{-1}\lambda(t)\,.$$

Hence we write a solution to equation (3.13) as:

$$(3.17)\qquad x(t) = (-DI - A')^{-1}\lambda(t)\,.$$

When (3.15) is only required to hold for $t \geq 0$, one may consider adding a term $\exp(-A't)\eta$ to solution (3.15) for x. In this case, however, any nonzero value of η results in a function x that is not in L_2^n no matter how this function is extended to the entire real line. Therefore, (3.15) gives the entire class of solutions in L_2^n.

4. Optimal Linear Regulator Problem

In this section we solve the following optimal linear regulator *(OLR)* problem.

OLR Problem: $$\max_{u\in L_+^N i,\, x\in L_+^n} -(1/2)\int_0^\infty [u(t)'\; x(t)'\; z(t)']\Omega \begin{bmatrix} u(t) \\ x(t) \\ z(t) \end{bmatrix} dt$$

subject to $x(t) = \exp(At)\mu + \int_0^t \exp(A\tau)\,[B_u u(t-\tau) + B_z z(t-\tau)]dt$ for $t \geq 0$.

Our formulation of the OLR problem differs from the standard formulation found, say, in Kwakernaak and Sivan (1972), in two ways. First, we allow z to be any element of $L_+^{n_z}$, not necessarily the solution to a linear differential equation system. Second, for systems in which it is not optimal to stabilize the state (i.e., for x to be in L_+^n), we impose stability as an extra constraint.

Without further restrictions, the OLR does not always have a solution. The following restriction on A, B_u and Ω guarantees the existence of a solution.

Assumption 4.1: The matrix $\Psi_{uu}(i\theta)$ is Hermitian, and there exists a strictly positive real number δ such that $\Psi_{uu}(i\theta) > \delta I$ for all θ in R, where

$$\Psi_{uu}(\zeta) \equiv [I\,;\; B_u'(-\zeta I - A')^{-1}]\Omega_{11} \begin{bmatrix} I \\ (\zeta I - A)^{-1}B_u \end{bmatrix} \tag{4.1}$$

and

$$\Omega_{11} = \begin{bmatrix} \Omega_{uu} & \Omega_{ux} \\ \Omega_{xu} & \Omega_{xx} \end{bmatrix} . \tag{4.2}$$

Note that for each θ in R, $\Psi_{uu}(i\theta)$ is a positive semidefinite matrix. Assumption 4.1 puts a lower bound on the eigenvalues of $\Psi(i\theta)$ for all θ. The inequality restriction must apply to the limit as θ goes to infinity as well. Evaluating this limit, we see that Assumption 4.1 implies that

$$\lim_{|\theta|\to\infty} \Psi_{uu}(i\theta) = \Omega_{uu} \geq \delta I \tag{4.3}$$

Therefore Ω_{uu} is nonsingular. In Appendix B we verify that when Assumptions 3.1 and 4.1 are satisfied, there exists a solution to the OLR problem for any initial condition μ.

We solve the OLR problem using Lagrange multipliers. Since L_+^n is a Hilbert space, continuous linear functionals on this space are representable as inner products with elements of L_+^n. Consequently, the Lagrange multiplier λ associated with constraint (3.12) is an element of L_+^n and the Lagrangean is given by:

(4.4)

$$\mathcal{L} \equiv -\int_0^\infty \left\{1/2[u(t)'\; x(t)'\; z(t)']\Omega \begin{bmatrix} u(t) \\ x(t) \\ z(t) \end{bmatrix} - \lambda(t)\cdot[x(t) - x^p(t) - x^h(t)]\right\}dt$$

It follows from the Parseval formula that

(4.5)

$$\begin{aligned}\int_0^\infty \lambda(t)\cdot x^p(t)\, dt &= (1/2\pi)\int_R T(\lambda)(\theta)^{*\prime}T(x^p)(\theta)\,d\theta \\ &= (1/2\pi)\int_R T(\lambda)(\theta)^{*\prime}(i\theta I - A)^{-1}BT(v)(\theta)\,d\theta \\ &= (1/2\pi)\int_R \{(-i\theta I - A')^{-1}T(\lambda)(\theta)\}^{*\prime}\,[B_u T(u)(\theta) \\ &\quad + B_z T(z)(\theta)]d\theta \\ &= \int_0^\infty \{[B_u'(-DI - A')^{-1}\lambda(t)\}\cdot u(t)dt \\ &\quad + \int_0^\infty \{[B_z'(-DI - A')^{-1}\lambda(t)\}\cdot z(t)dt .\end{aligned}$$

The first-order conditions for Problem 1 are obtained by differentiating $\mathcal{L}$ with respect to u and x, giving

$$\Omega_{uu}\, u_s(t) + \Omega_{ux}\, x_s(t) + \Omega_{uz}\, z(t) + B_u'\,(-DI - A')^{-1}\,\lambda_s(t) = 0 \tag{4.6}$$

$$\Omega_{xu}\, u_s(t) + \Omega_{xx}\, x_s(t) + \Omega_{xz}\, z(t) = \lambda_s(t) . \tag{4.7}$$

The subscripts s are included on x, u, and λ to denote the solution.

We solve equation (4.6) for $u_s(t)$ giving

$$\begin{aligned} u_s(t) = &-(\Omega_{uu})^{-1}B_u'(-DI - A')^{-1}\lambda_s(t) - (\Omega_{uu})^{-1}\Omega_{ux}\,x_s(t) \\ &-(\Omega_{uu})^{-1}\Omega_{uz}\,z(t) .\end{aligned} \tag{4.8}$$

Substituting (4.8) into (4.7) yields

$$\begin{aligned}\lambda_s(t) = &-\Omega_{xu}\,(\Omega_{uu})^{-1}B_u'\,(-DI - A')^{-1}\lambda_s(t) \\ &+[\Omega_{xx} - \Omega_{xu}\,(\Omega_{uu})^{-1}\Omega_{ux}]x_s(t) \\ &+[\Omega_{xz} - \Omega_{xu}\,(\Omega_{uu})^{-1}\Omega_{uz}]z(t) .\end{aligned} \tag{4.9}$$

For each time $t \geq 0$, define a co-state vector $x_c(t)$ via

$$x_c(t) \equiv (-DI - A')^{-1} \lambda_s(t) \ . \tag{4.10}$$

Then the analysis in Section 3 implies that the co-state equation

$$Dx_c(t) = -A'x_c(t) - \lambda_s(t) \tag{4.11}$$

is satisfied [see (3.13) - (3.17)]. Substituting (4.9) and (4.10) into (4.11) gives

(4.12)
$$\begin{aligned} Dx_c(t) = [\Omega_{xu}\,(\Omega_{uu})^{-1}\,\Omega_{ux} - \Omega_{xx}]x_s(t) - [A - B_u(\Omega_{uu})^{-1}\Omega_{ux}]'x_c(t) \\ + [\Omega_{xu}\,(\Omega_{uu})^{-1}\Omega_{uz} - \Omega_{xz}]z(t) \ . \end{aligned}$$

Substituting (4.8) and (4.10) into the state equation (1.13) yields

$$\begin{aligned} Dx_s(t) = [A - B_u(\Omega_{uu})^{-1}\Omega_{ux}]x_s(t) - B_u(\Omega_{uu})^{-1}B_u'x_c(t) \\ + [B_z - B_u(\Omega_{uu})^{-1}\Omega_{uz}]z(t) \ . \end{aligned} \tag{4.13}$$

To solve for $x_s(t)$ and $x_c(t)$, we combine equations (4.12) and (4.13) into a single system:

$$\begin{bmatrix} Dx_s(t) \\ Dx_c(t) \end{bmatrix} = H \begin{bmatrix} x_s(t) \\ x_c(t) \end{bmatrix} + Kz(t) \ . \tag{4.14}$$

where

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \quad K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}, \tag{4.15}$$

$H_{11} \equiv A - B_u(\Omega_{uu})^{-1}\Omega_{ux}$, $H_{12} \equiv -B_u(\Omega_{uu})^{-1}B_u'$, $H_{21} \equiv -\Omega_{xx} + \Omega_{xu}(\Omega_{uu})^{-1}\Omega_{ux}$ and $H_{22} \equiv -[A - B_u(\Omega_{uu})^{-1}\Omega_{ux}]'$, $K_1 \equiv B_z - B_u(\Omega_{uu})^{-1}\Omega_{uz}$ and $K_2 \equiv \Omega_{xu}(\Omega_{uu})^{-1}\Omega_{uz} - \Omega_{xz}$. The matrix H is referred to as a Hamiltonian matrix and satisfies two properties that are important for our purposes. First $H_{22} = -H_{11}'$; and second, H_{12} and H_{21} are symmetric.

Consider first the case in which $z(t)$ is zero for all t. In this case we solve the homogeneous differential equation system:

$$\begin{bmatrix} Dx_s(t) \\ Dx_c(t) \end{bmatrix} = H \begin{bmatrix} x_s(t) \\ x_c(t) \end{bmatrix} \tag{4.16}$$

We follow Vaughan (1969) and solve (4.16) by taking the Jordan decomposition of H:

$$H = EJ(E)^{-1} \tag{4.17}$$

where J is a matrix containing the eigenvalues of H on the diagonal. Some of the entries of J immediately to the right of the eigenvalues that occur more than once are one, and the remaining entries of J are zero. For notational convenience, we place the eigenvalues with strictly negative real parts in the upper left block of J.

The eigenvalues of H occur in symmetric pairs *vis-a-vis* the imaginary axis of the complex plane. To see this, let ζ be any eigenvalue of H, and let $[e_1', \ e_2']'$ be the corresponding column eigenvector of H. Then $[e_2', \ -e_1']$ is a row eigenvector of H with eigenvalue $-\zeta$. As a consequence of this symmetry, there can only be at most n eigenvalues of H with strictly negative real parts. We will show, in fact, that there are exactly n such eigenvalues.

It is convenient to transform the equation system (4.14). Define

$$x^+(t) = E^{-1} \begin{bmatrix} x_s(t) \\ x_c(t) \end{bmatrix} \tag{4.18}$$

Then $x^+(t)$ satisfies

$$Dx^+(t) = Jx^+(t) \ . \tag{4.19}$$

The advantage of working with differential equation system (4.19) is the dynamics are uncoupled according to the distinct eigenvalues.

We partition $x^+(t)' = [x_n^+(t)', \ x_p^+(t)']$ where the number of entries of $x_n^+(t)$ is equal to the number of eigenvalues of H with strictly negative real parts. Since the remaining eigenvalues have nonnegative real parts, the function x_p^+ will not be stable (have entries in L_+^1) unless $x_p^+(0)$ is zero. Thus we have two initial conditions: $x_s(0) = \mu$ and $x_p^+(0) = 0$.

It follows from (4.18) that

$$E \begin{bmatrix} x_n^+(t) \\ 0 \end{bmatrix} = \begin{bmatrix} x_s(t) \\ x_c(t) \end{bmatrix} \ . \tag{4.20}$$

Partition E as

$$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} \ . \tag{4.21}$$

Then for $t = 0$, (4.20) can be expressed as

$$E_{11}\, x_n^+(0) = \mu \quad \text{and} \quad E_{21}\, x_n^+(0) = x_c(0) \tag{4.22}$$

Since a solution is known to exist (see Appendix B), for any μ there must exist a vector $x_n^+(0)$ such that (4.22) is satisfied. Consequently, E_{11} must have n columns (H must have n eigenvalues with strictly negative real parts) and E_{11} must be nonsingular. Solving (4.22) for $x_c(0)$ gives

$$x_c(0) = M_x\, \mu\ , \tag{4.23}$$

where $M_x \equiv E_{21}\,(E_{11})^{-1}$ and solving (4.20) for $x_c(t)$ gives

$$x_c(t) = M_x x_s(t)\ . \tag{4.24}$$

It is of interest to obtain recursive solutions for $Dx_s(t)$ and $Dx_c(t)$. Substituting (4.24) into (4.18) gives

$$\begin{bmatrix} Dx_s(t) \\ Dx_c(t) \end{bmatrix} = \begin{bmatrix} (H_{11} + H_{12}M_x) \\ (H_{21} - H_{11}'M_x) \end{bmatrix} x_s(t) \tag{4.25}$$

The eigenvalues of $H_{11} + H_{12}M_x$ are the same as the stable eigenvalues of H. It follows from (4.24) that $Dx_c(t) = M_x Dx_s(t)$ and from (4.25) that

$$-M_x H_{11} - M_x H_{12} M_x + H_{21} - H_{11}' M_x = 0\ . \tag{4.26}$$

An equivalent statement of (4.26) is

$$[-M_x\ I]\, H \begin{bmatrix} I \\ M_x \end{bmatrix} = 0\ . \tag{4.27}$$

In Appendix C we show that the matrix M_x turns out to be symmetric and positive semidefinite. Furthermore, the time zero value function, or equivalently the optimal value of the criterion function as a function of the initial condition for $x(t)$, is given by $(-1/2)\mu' M_x \mu$.

We now consider the general case in which $z(t)$ is allowed to be different from zero. In this case it is convenient to use a somewhat different transformation than (4.18). Define a matrix M

$$M \equiv \begin{bmatrix} I & 0 \\ -M_x & I \end{bmatrix} \tag{4.28}$$

and transform the composite state-costate function:

$$\begin{bmatrix} x_s(t) \\ w(t) \end{bmatrix} \equiv M \begin{bmatrix} x_s(t) \\ x_c(t) \end{bmatrix}\ . \tag{4.29}$$

Then

$$\begin{bmatrix} Dx_s(t) \\ Dw(t) \end{bmatrix} = MHM^{-1} \begin{bmatrix} x_s(t) \\ w(t) \end{bmatrix} + MKz(t) \tag{4.30}$$

Since M_x is symmetric and (4.27) is satisfied,

$$MHM^{-1} = \begin{bmatrix} (H_{11} + H_{12}M_x) & H_{12} \\ 0 & -(H_{11} + H_{12}M_x)' \end{bmatrix} \tag{4.31}$$

and

$$MK = \begin{bmatrix} K_1 \\ -M_x K_1 + K_2 \end{bmatrix}\ . \tag{4.32}$$

It is possible to solve the second block of equations in (4.30) separately because the matrix $MH(M)^{-1}$ is upper triangular. The eigenvalues of the matrix $-(H_{11}+H_{12}M_x)'$ have strictly positive real parts: therefore, the stable solution for $w(t)$ is given by the following forward convolution:

$$\begin{aligned} w(t) &= [-DI - (H_{11} + H_{12}M_x)']^{-1}\,(K_2 - M_x K_1) z(t) \\ &= \int_0^\infty \exp[(H_{11} + H_{12}M_x)'\tau]\ (K_2 - M_x K_1) z(t+\tau) d\tau \end{aligned} \tag{4.33}$$

It follows from (4.30) that

$$Dx_s(t) = (H_{11} + H_{12}\, M_x) x_s(t) + H_{12}\, w(t) + K_1\, z(t)\ . \tag{4.34}$$

We refer to $(H_{11} + H_{12}\, M_x)x_s(t)$ as the *feedback* part and $H_{12}\, w(t) + K_1\, z(t)$ as the *feedforward* part of the decision rule for $Dx_s(t)$. Notice that the transpose of the feedback matrix $(H_{11} + H_{12}\, M_x)$ enters the exponential in the feedforward integral.

5. Recursive Solution Methods

In this section we describe recursive methods for computing two objects. First we show how to calculate the matrix M_x used in the feedback and feedforward portions of the decision rule for $Dx_s(t)$. Then we show how to calculate a solution for $w(t)$ as a function of $z(t)$ in the special case in which $z(t)$ is the solution to a first-order differential equation system.

Consider first the matrix M_x. When $z(t)$ is set to zero for all t, the constraints are autonomous differential equations. Roberts (1971) and Denman and Beavers (1976) proposed a matrix sign algorithm for solving the resulting optimization problem and in particular for computing M_x. An attractive feature of the matrix sign algorithm is that it avoids computing the Jordan Decomposition of the Hamiltonian matrix H. Instead the matrix M_x is computed by calculating successively the average of a matrix and its inverse. The algorithm is initialized at the matrix H. More precisely, for any nonsingular matrix G, define $\mathcal{R}(G)$ via:

(5.1) $$\mathcal{R}(G) = (1/2)\ [G + (G)^{-1}]\ .$$

Consider the sequence $\{\mathcal{R}^j(H)\}$ where $\mathcal{R}^j$ denotes the mapping $\mathcal{R}$ applied j times in succession. Using the Jordan decomposition for H, it follows that

(5.2) $$\mathcal{R}^j(H) = E\mathcal{R}^j(J)E^{-1}\ .$$

Recall that J is a matrix with the eigenvalues of H on the diagonal. The first n diagonal entries contain the eigenvalues with strictly negative real parts and the remaining n diagonal entries contain the eigenvalues with strictly positive real parts. Some of the entries of J immediately to the right of eigenvalues that occur more than once are one and the remaining entries of J are zero. As a result, the sequence $\{\mathcal{R}^j(J)\}$ converges to a diagonal matrix with minus ones for the first n entries and ones for the second n entries.

Let H^∞ denote the limit point of $\{\mathcal{R}^j(H)\}$ and partition this matrix:

(5.3) $$H^\infty = \begin{bmatrix} H_{11}^\infty & H_{12}^\infty \\ H_{21}^\infty & H_{22}^\infty \end{bmatrix}\ .$$

Recall that E is partitioned as

(5.4) $$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}\ .$$

Similarly, partition $(E)^{-1}$ as

(5.5) $$(E)^{-1} = \begin{bmatrix} E^{11} & E^{12} \\ E^{21} & E^{22} \end{bmatrix}\ .$$

Then

(5.6) $$H_{11}^\infty = -E_{11}\,E^{11} + E_{12}\,E^{21} = -2E_{11}\,E^{11} + I\ ,$$

and

(5.7) $$H_{21}^\infty = -E_{21}\,E^{11} + E_{22}\,E^{21} = -2E_{21}\,E^{11}\ .$$

As long as E^{11} is nonsingular,

(5.8) $$M_x = H_{21}^\infty\,(H_{11}^\infty - I)^{-1}\ .$$

Hence M_x can be approximated by computing $\mathcal{R}^j(H)$ for a sufficiently large value of j and then applying formula (5.8).

The restriction that E^{11} be nonsingular is not satisfied for some interesting parameter configurations of our model. For instance, E^{11} can be singular when it is not optimal for the state function x to be in L_+^n. When E^{11} is singular, we add a quadratic penalty $\hat{\delta}I$ to Ω_{xx} in the criterion for the OLR problem where $\hat{\delta}$ is strictly positive. This ensures that it is optimal for x to be in L_+^n. By selecting $\hat{\delta}$ to be sufficiently small, we can approximate the matrix M_x with arbitrary accuracy.

Anderson (1978) showed how to exploit the structure of the Hamiltonian matrix to simplify the iterations in the matrix sign algorithm. Let G be any Hamiltonian matrix. Anderson (1978) suggested the following partitioning:

(5.9) $$\mathcal{R}(G)_{11} = (1/2)\ \{G_{11} + [G_{11} + G_{12}(G_{11}')^{-1}G_{21}]^{-1}\}$$

(5.10) $$\mathcal{R}(G)_{12} = (1/2)\ \{G_{11} + [G_{11} + G_{12}(G_{11}')^{-1}G_{21}]^{-1}G_{12}\ (G_{11}')^{-1}\}$$

(5.11) $$\mathcal{R}(G)_{21} = (1/2)\ \{G_{12} + (G_{11}')^{-1}G_{21}[G_{11} + G_{12}(G_{11}')^{-1}\ G_{21}]^{-1}\}$$

where $\mathcal{R}(G)_{11}$ is the upper right n by n block of $\mathcal{R}(G)$, and so on. This partitioning exploits the fact that the (2,2) block of the Hamiltonian matrix is the negative transpose of the (1,1) block. The matrix sign algorithm preserves this relation among the partitions of a matrix.

Consider next the computation of $w(t)$. For general specifications of $z(t)$, it is not possible to simplify further (4.33). We now impose some additional structure on $z(t)$. Suppose that $z(t)$ is the solution to the differential equation system:

(5.12) $$Dz(t) = A_{zz}\,z(t)$$

where the eigenvalues of A_{zz} have real parts that are strictly negative. In this case, $z(t+\tau)$ is $\exp(A_{zz}\,\tau)z(t)$ implying that

$$w(t) = M_z\, z(t) \tag{5.13}$$

where

$$M_z = \int_0^\infty \exp[(H_{12}\, M_x + H_{11})'\tau]\,(K_2 - M_x\, K_1)\ \exp(A_{zz}\,\tau)d\tau\ . \tag{5.14}$$

There are two alternative ways of computing M_z. Using integration-by-parts, it can be shown that M_z satisfies:

$$(H_{12}\, M_x + H_{11})'\, M_z + M_z\, A_{zz} = K_2 - M_x\, K_1 \tag{5.15}$$

This equation is linear in the entries of the matrix M_z. Furthermore, since the eigenvalues of $(H_{12}\, M_x + H_{11})'$ and A_{zz} have strictly negative real parts, it can be shown that M_z is the unique solution to (5.15). Therefore, one way to compute M_z is to write (5.15) as a system of linear equations in the entries of M_z and solve that system of equations.

Following Denman and Beavers (1976), a second approach is to note that $w(t)$ and $z(t)$ satisfy the composite first-order differential equation system:

$$\begin{bmatrix} Dw(t) \\ Dz(t) \end{bmatrix} = \hat{H} \begin{bmatrix} w(t) \\ z(t) \end{bmatrix} \tag{5.16}$$

where

$$\hat{H} = \begin{bmatrix} -(H_{12}\, M_x + H_{11})' & (K_2 - M_x\, K_1) \\ 0 & A_{zz} \end{bmatrix}, \tag{5.17}$$

and to apply the matrix sign algorithm to find the stable solution to the composite system. Since $\hat{H}$ is upper triangular, the collection of eigenvalues of $\hat{H}$ are given by the union of the collection of eigenvalues of $-(H_{12}\, M_x + H_{11})'$ and the collection of eigenvalues of A_{zz}. The matrix M_z is found by mimicking the approach used to solve differential equation system (4.16). In particular, we deduce the limit point to the sequence $\{\mathcal{R}^j(\hat{H})\}$ and use the analog to formula (5.8) to compute M_z.

The matrix $\hat{H}$ is upper block triangular, and the matrix sign algorithm preserves this triangularity. These features can be exploited by partitioning the algorithm. Let $\hat{G}$ be an upper triangular matrix partitioned as

$$\hat{G} = \begin{bmatrix} \hat{G}_{11} & \hat{G}_{12} \\ 0 & \hat{G}_{22} \end{bmatrix}. \tag{5.18}$$

Then the matrix sign algorithm can be partitioned as:

$$\mathcal{R}(\hat{G})_{11} = (1/2)\ [\hat{G}_{11} + (\hat{G}_{11})^{-1}] \tag{5.19}$$

$$\mathcal{R}(\hat{G})_{12} = (1/2)\ [\hat{G}_{12} - (\hat{G}_{11})^{-1}\,\hat{G}_{12}\,(\hat{G}_{22})^{-1}] \tag{5.20}$$

$$\mathcal{R}(\hat{G})_{22} = (1/2)\ [\hat{G}_{22} + (\hat{G}_{22})^{-1}] \tag{5.21}$$

with $\mathcal{R}(\hat{G})_{21} = 0$. Finally, since the upper left block of $\hat{H}$ has eigenvalues with strictly positive real parts and the lower right block has eigenvalues with strictly negative real parts, it is straightforward to show that $\mathcal{R}^\infty(\hat{H})_{11} = I$ and that $\mathcal{R}^\infty(\hat{H})_{22} = -I$. The analog to (5.8) is simply

$$M_z = (-1/2)\mathcal{R}^\infty\,(\hat{H})_{12}\ . \tag{5.22}$$

When $z(t)$ satisfies (5.12), it is possible to augment the state vector $x(t)$ to include $z(t)$. An alternative solution approach is to form the Hamiltonian matrix for this augmented system and to initialize the matrix sign algorithm at this matrix. Although the approach suggested in this section requires two applications of the matrix sign algorithm, in each stage the algorithm is applied to matrices with smaller dimensions. As a consequence, the dual applications of the matrix sign algorithm typically will be faster than the single application to the augmented Hamiltonian matrix. This two-stage algorithm is the continuous time counterpart to a two-stage algorithm for solving discrete time quadratic control problems suggested by Hansen and Sargent (1981a).

6. Adjustment Cost Model of Investment

In this section we investigate the solution to the *OLR* problem in a special set of circumstances. We focus on competitive equilibrium models of investment under rational expectations as in Lucas (1981), Lucas and Prescott (1971) and Brock and MaGill (1979). The technologies we consider are the continuous time counterparts to technologies studied in Hansen and Sargent (1981a). There is a vector of capital stocks and there are costs to adjusting these stocks as in Lucas (1967), Gould (1968), Treadway (1969) and Mortenson (1973). The adjustment costs may be of higher order than one, which in the context of continuous time models may apply to higher derivatives of the capital stocks. The economic environment is described in Subsection 6.A. In Subsection 6.B we derive an alternative representation of the first-order conditions that can be interpreted as Euler equations. In Subsection 6.C we describe an alternative approach to solving the *OLR* problem and relate this approach to the one described in Section 5.

6a. Setup

Consider a special case of the general model presented in Section 1. Suppose there is no household capital so that (1.2) is replaced by

$$s(t) = \Pi c(t) \ . \tag{6.1}$$

Let $\hat{y}(t)$ be an m-dimensional vector of productive capital stocks at time t, and define $\hat{x}(t)' = k(t)' \equiv [D^{\ell-1}\hat{y}(t)', \ D^{\ell-2}\hat{y}(t)', \ldots, \hat{y}(t)']$ and $\hat{u}(t) \equiv D^{\ell}\hat{y}(t)$. In this case,

$$D\hat{x}(t) = \hat{A}\hat{x}(t) + B_u \hat{u}(t) \tag{6.2}$$

where

$$\hat{A} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ I & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & I & 0 \end{bmatrix} \quad B_u = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \tag{6.3}$$

Resource constraint (1.5) is imposed as before with $\hat{u}(t) = i(t)$. Notice that depreciation in capital is not reflected in specification of $\hat{A}$. This is because we have let the control be a measure of net investment. Suppose instead that $i(t)$ denotes a measure of gross investment and that

$$i(t) = \Gamma^* k(t) + \hat{u}(t) \tag{6.4}$$

where $\Gamma^* k(t)$ measures the investment required to replace any depreciated capital stock. Then (1.5) and (6.4) imply that

$$\Phi_c c(t) + \Phi_i \hat{u}(t) = (\Gamma - \Phi_i \Gamma^*)k(t) + f(t) \ . \tag{6.5}$$

Hence our use of net investment instead of gross investment as the control function requires that we replace Γ by $(\Gamma - \Phi_i \Gamma^*)$ in constraint (1.5).

6b. Euler Equations

Let $y(t) = \exp(-\epsilon t)\hat{y}(t)$. It is convenient to represent $x(t)$ in terms of $y(t)$ and its derivatives. Note that

$$Dy(t) = -\epsilon y(t) + \exp(-\epsilon t)D\hat{y}(t) \ . \tag{6.6}$$

In operator notation we have that

$$(D+\epsilon)y(t) = \exp(-\epsilon t)D\hat{y}(t) \ . \tag{6.7}$$

Similarly,

$$(D+\epsilon)^j y(t) = \exp(-\epsilon t)D^j\hat{y}(t) \tag{6.8}$$

where the operator raised to a power denotes sequential application of the operator. Hence we can represent $x(t)$ as:

$$x(t) = \begin{bmatrix} (D+\epsilon)^{\ell-1}I \\ (D+\epsilon)^{\ell-2}I \\ \vdots \\ I \end{bmatrix} y(t) \ . \tag{6.9}$$

We use this structure to obtain an alternative representation of the first-order conditions (4.6) and (4.7). Define a matrix function P of a complex variable:

$$P(\zeta) \equiv \begin{bmatrix} (\zeta+\epsilon)^{\ell}I \\ (\zeta+\epsilon)^{\ell-1}I \\ \vdots \\ I \end{bmatrix} , \tag{6.10}$$

First-order condition (4.6) can be expressed as

$$[\Omega_{uu}\,\Omega_{ux}]\,P(D)y(t) + \Omega_{uz}\,z(t) = -B_u'(-DI - A')^{-1}\lambda_s(t) \ . \tag{6.11}$$

Equations of system (4.7) can be expressed as

$$[\Omega_{xu}\,\Omega_{xx}]\,P(D)y(t) + \Omega_{xz}\,z(t) = \lambda_s(t) \ . \tag{6.12}$$

Recall that A' is given by

$$A' = \begin{bmatrix} -\epsilon I & I & 0 & \ldots & 0 & 0 \\ 0 & -\epsilon I & I & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & -\epsilon I \end{bmatrix} \tag{6.13}$$

and that $B_u' = [I\ 0]$. The matrix operator $B_u'(-\zeta I - A')^{-1}$ consists of the first m rows of the operator $(-\zeta I - A')^{-1}$. Then

(6.14)
$$-B_u'(-\zeta I - A')^{-1} = -[(-\zeta+\epsilon)^{-1}I;\ (-\zeta+\epsilon)^{-2}I;\ \ldots\ (-\zeta+\epsilon)^{-\ell}I] \ .$$

Substituting into (6.11) we obtain

$$(6.15)\qquad \begin{aligned}&[\Omega_{uu}\,\Omega_{ux}]\,P(D)y(t)+\Omega_{uz}\,z(t)+\\ &\big[(-D+\epsilon)^{-1}I\ ;\ (-D+\epsilon)^{-2}I\ ;\ \ldots\ (-D+\epsilon)^{-\ell}I\big]\\ &\{[\Omega_{xu}\,\Omega_{xx}]\,P(D)y(t)+\Omega_{xz}\,z(t)\}=0\ .\end{aligned}$$

Define a matrix function Q of a complex variable ζ:

$$(6.16)\qquad \begin{aligned}Q(\zeta)&\equiv[I\ ;\ (-\zeta+\epsilon)^{-1}I\ ;\ (-\zeta+\epsilon)^{-2}I\ ;\ \ldots\ (-\zeta+\epsilon)^{-\ell}I]\\ &=(-\zeta+\epsilon)^{-\ell}\,P(-\zeta)'\ .\end{aligned}$$

Then (6.15) can be expressed as

$$(6.17)\qquad [Q(D)\Omega_{11}P(D)]y(t)=-Q(D)\Omega_{12}\,z(t)$$

where $\Omega_{11}\ \equiv\ \begin{bmatrix}\Omega_{uu}&\Omega_{ux}\\ \Omega_{xu}&\Omega_{xx}\end{bmatrix}$ and $\Omega_{12}=\begin{bmatrix}\Omega_{uz}\\ \Omega_{xz}\end{bmatrix}$.

In (6.17) $Q(D)$ is a *forward* convolution operator and $P(D)$ a *backwards* derivative operator.

Finally, it is of interest to disentangle the effect of scaling by $\exp(-\epsilon t)$. Hence we deduce a corresponding equation for $\hat{y}(t)$ in terms of $\hat{z}(t)$. It follows from (6.6) that

$$(6.18)\qquad Dy(t)=\exp(-\epsilon t)\ [(D-\epsilon)\hat{y}(t)]\ .$$

To undo the scaling in (6.17), we simply substitute $(D-\epsilon)$ for D and multiply by $\exp(\epsilon t)$:

$$(6.19)\qquad [Q(D-\epsilon)\Omega_{11}\,P(D-\epsilon)]\,\hat{y}(t)=-Q(D-\epsilon)\,\Omega_{12}\,\hat{z}(t)\ .$$

These are the Euler equations for the optimization problem.

6c. Solution

We now describe an alternative Euler equation approach to solving equations (6.19). In light of (6.16)

$$(6.20)\qquad Q(\zeta)\,\Omega_{11}\,P(\zeta)=P(-\zeta)'\,\Omega_{11}\,P(\zeta)/(-\zeta+\epsilon)^{\ell}\ .$$

Divide both sides of (6.20) by $(\zeta+\epsilon)^{\ell+1}\ (-\zeta+\epsilon)$ and evaluate the resulting function at $\zeta=i\theta$:

$$(6.21)\qquad \begin{aligned}F(\theta)&\equiv P(-i\theta)'\,\Omega_{11}P(i\theta)/[(-i\theta+\epsilon)\,(i\theta+\epsilon)]^{\ell+1}\\ &=\Psi_{uu}(i\theta)/[(-i\theta+\epsilon)\,(i\theta+\epsilon)]\end{aligned}$$

By virtue of Assumption 3.1, $F(\theta)$ is a positive definite matrix for all θ in R. In addition, F is a rational function of θ, $F(-\theta)=F(\theta)'$ and

$$(6.22)\qquad \lim_{\theta\to\infty}F(\theta)=0\ .$$

As a consequence, F is the spectral density function for a linearly regular, stochastically nonsingular continuous time stochastic process. Since the spectral density function is rational, with common denominator $[(-i\theta+\epsilon)\,(i\theta+\epsilon)]^{\ell+1}$, it can be factored:

$$(6.23)\qquad F(\theta)=\hat{P}(-i\theta)'V\hat{P}(i\theta)/[(-i\theta+\epsilon)\,(i\theta+\epsilon)]^{\ell+1}$$

where V is a nonsingular symmetric matrix, $\hat{P}(\zeta)$ is a polynomial of degree ℓ that is nonsingular in the left-half plane of C and

$$(6.24)\qquad \lim_{\zeta\to\infty}\hat{P}(\zeta)/\zeta^{\ell}=I$$

(e.g. see Rozanov 1967). As a consequence, the composite operator $[Q(D)\Omega_{11}P(D)]$ used in Euler equation (6.16) can be factored

$$(6.25)\qquad Q(D)\Omega_{11}P(D)=[\hat{P}(-D)/(-D+\epsilon)^{\ell}]'\,V\hat{P}(D)\ .$$

The operator $\hat{P}(D)$ is a backwards derivative operator and the operator $[\hat{P}(-D)/(-D+\epsilon)^{\ell}]$ is a forward convolution that can be expressed as the matrix linear combination of the identity operator and the forward convolution operators $1/(-D+\epsilon)$, $1/(-D+\epsilon)^2,\ldots 1/(-D+\epsilon)^{\ell}$.

The factorization given in (6.25) is of interest because the rational function $[\hat{P}(-\zeta)/(-\zeta+\epsilon)^{\ell}]$ is nonsingular in the left-half plane of C. As a consequence the forward operator $[\hat{P}(-D)/(-D+\epsilon)^{\ell}]$ has a one-sided forward inverse that can be characterized by inverting $[\hat{P}(-\zeta)/(-\zeta+\epsilon)^{\ell}]$. Let

$$(6.26)\qquad \hat{Q}(\zeta)\equiv V^{-1}(-\zeta+\epsilon)^{\ell}\,\hat{P}(-\zeta)'^{-1}\,Q(\zeta)\Omega_{12}\ .$$

Then the entries of $\hat{Q}(\zeta)$ are rational functions of ζ with numerator orders that do not exceed the denominator orders. In addition, the poles of $\hat{Q}(\zeta)$ reside in the left-half plane of C. As a consequence $\hat{Q}(D)$ can be expressed as a matrix linear combination of the identity operator and convolution operators. The terms of this matrix linear combination can be deduced by computing a matrix partial fractions decomposition of $\hat{Q}(\zeta)$.

When Π is strictly positive, the original formulation of the model still applies. In this case current consumption and the household capital stock are substitutes in the production of consumption services. Although the consumption good is durable, additional services are generated by new acquisitions of the goods.

Habit persistence in preferences over consumption as examined for example by Pollak (1970), Ryder and Heal (1973), Sundaresan (1989), Constantinides (1990) and Heaton (1989) can be accommodated by assuming that $\Lambda = -1$, $\Pi \geq 1$. In this case consumption and capital are complements in the production of consumption services.[1]

When $\Lambda = 0$ and $\Pi = 1$, the induced preferences for consumptions are time separable. The solution method described in Section 6 can be applied to characterize the optimal law of motion for capital. In this case, we can drop the household capital stock, and the productive capital stock becomes the only component of the endogenous state variable. The restriction that $k(t)$ discounted by $\exp(-\epsilon t)$ be in L^1_+ is important in obtaining a solution that is of interest. Without this restriction, $c(t)$ is set to $b(t)$ which in turn is supported by an unstable time path for the discounted capital stock.

To solve this model we let $z(t)' = [b(t), f(t)]$. The matrices Ω_{11} and Ω_{12} of Section 6 are given by:

$$(7.3) \qquad \Omega_{11} = \begin{bmatrix} 1 & -\rho \\ -\rho & \rho^2 \end{bmatrix} \quad \text{and} \quad \Omega_{12} = \begin{bmatrix} -1 \\ \rho \end{bmatrix} [-1,\, 1] \,.$$

The composite operator $Q(D)\Omega_{11}P(D)$ in this case is given by:
(7.4)

$$[1 \;\; 1/(-D+\epsilon)] \begin{bmatrix} 1 & -\rho \\ -\rho & \rho^2 \end{bmatrix} \begin{bmatrix} (D+\epsilon) \\ 1 \end{bmatrix} = (-D+\epsilon)\,(D+\epsilon)/(-D+\epsilon) \,.$$

Hence the operator $\hat{P}(D)$ of equation (6.23) is $(D+\epsilon)$ and V is equal to 1. The operator $-Q(D)\Omega_{12}$ is given by:

$$(7.5) \qquad \begin{aligned} -Q(D)\Omega_{12} &= -[1 \;\; 1/(-D+\epsilon)] \begin{bmatrix} -1 \\ \rho \end{bmatrix} [-1,\, 1] \\ &= [1 - \rho/(-D+\epsilon)]\, [-1,\, 1] \,. \end{aligned}$$

The operator $\hat{Q}(D)$ of equation (6.26) is

$$(7.6) \qquad \begin{aligned} \hat{Q}(D) &= (-D+\epsilon)\, [1 - \rho/(-D+\epsilon)]\, [-1,\, 1]/(-D+\epsilon) \\ &= [1 - \rho/(-D+\epsilon)]\, [-1,\, 1] \,. \end{aligned}$$

It follows from (6.28) that the capital stock satisfies:

(7.7)

$$\begin{aligned} Dk(t) &= [1 - \rho/(-D+\epsilon)]\, [-1,\, 1]\, z(t) \\ &= [1 - \rho/(-D+\epsilon)]\, [f(t) - b(t)] \\ &= [f(t) - b(t)] - \rho \int_0^\infty \exp(-\rho\tau) - [f(t+\tau) - b(t+\tau)] d\tau \,. \end{aligned}$$

In other words, investment is calculated by comparing the current level of the endowment relative to the satiation point to a weighted average of future endowments relative to satiation points. Heaton (1989) displays solutions to the model for other settings for Λ and Π.

Appendix A

In this appendix we verify that formulas (2.3), (2.6), (2.8), and (2.15) apply to all of L_2^n. Let 1_τ denote the indicator function of the set $[-\tau, \tau]$, and let x denote any member of L_2^n. Then $1_\tau x$ is in $L_1^n \cap L_2^n$ for each t and $\{1_\tau x : t \geq 1\}$ converges in L_2^n to x. Since T is continuous, $\{T(1_\tau x) : t \geq 1\}$ converges in L_2^n to $T(x)$. Let O be any of the four operators introduced in Section 2, and let ϕ be the corresponding function such that $\phi T(x) = T[O(x)]$. The function ϕ is bounded in all four cases. Therefore, $\{T[O(1_\tau x)] : \tau \geq 1\}$ converges in L_2^n to $T[O(x)]$. The Parseval formula implies that $\{O(1_\tau x) : \tau \geq 1\}$ converges in L_2^n to $O(x)$. A subsequence $\{O[1_{\tau(j)} x] : j \geq 1\}$ converges pointwise to $O(x)$ except on a set of measure zero. Formulas (2.3), (2.6), (2.8) and (2.15) involve integral representations. In this case, $O(1_\tau x)$ has an integral representation. Since x can be expressed as a linear combination of nonnegative, vectors of real-valued functions, it follows from the Monotone Convergence Theorem that $O(x)$ has the same integral representation.

Appendix B

In this appendix we establish that there exists a solution to the *OLR* problem when Assumption 4.1 is satisfied. Since the matrix Ω is positive semidefinite, the criterion function is always less than or equal to zero. Define

$$(B.1) \qquad T(u,\, x) \equiv \int_0^\infty [u(t)' \; x(t)' \; z(t)']\Omega \begin{bmatrix} u(t) \\ x(t) \\ z(t) \end{bmatrix} dt \,,$$

and put

$$(B.2) \quad \bar{\delta} \equiv \inf\{T(u,\, x) : u \in L_+^m,\; x \in L_+^n,\; (u,\, x) \text{ satisfies } (1.13)\} \,.$$

Notes

1. For some values of $\Pi < 1$, the household technology generates *rational addiction* as suggested by Stigler and Becker (1977) and Becker and Murphy (1988).

8

Prediction Formulas for Continuous Time Linear Rational Expectations Models

by Lars Peter HANSEN and Thomas J. SARGENT

In this note we derive optimal prediction formulas to be used in solving continuous time rational expectations models. In these derivations we employ Laplace transforms in a manner analogous to the use of z transforms for solving discrete time optimal prediction problems in Hansen and Sargent (1980a, Appendix A). The formulas are intended to play the same role for continuous time models that the discrete time formulas for optimal predictions of geometric distributed leads did in Hansen and Sargent (1980a).

1. Convolutions and Prediction

Let L^1 and L^2 denote the spaces of all real-valued Borel measurable functions ϕ on $\mathbf{R}$ that are absolutely integrable and square integrable, respectively. Let W denote a random measure defined on $\mathbf{R}$ with increments that are orthogonal and second-moment stationary. In other words,

$$E\left[W\{[t_2,\ t_1)\}^2\right] = t_2 - t_1 \quad \text{for} \quad t_2 > t_1 \ , \tag{1.1}$$

and

$$E\left[W\{[t_4,\ t_3)\}W\{[t_2,\ t_1)\}\right] = 0 \quad \text{for} \quad t_4 > t_3 > t_2 > t_1 \ . \tag{1.2}$$

Using functions in L^2 and the random measure W, we construct second-moment stationary processes as convolutions:

$$x(t) = \int_{-\infty}^{+\infty} \phi(\tau) dW(t-\tau) \ . \tag{1.3}$$