



TI 2001-043/2

Tinbergen Institute Discussion Paper

Reducing the Dimensionality of Linear Quadratic Control Problems

Ronald J. Balvers^{1,2}

Douglas W. Mitchell¹

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>

Reducing the Dimensionality of Linear Quadratic Control Problems*

February 2001

Ronald J. Balvers
Department of Economics
P.O. Box 6025
West Virginia University
Morgantown, WV 26506-6025
Phone: (304) 293-7880
Email: rbalvers@wvu.edu

Douglas W. Mitchell
Department of Economics
P.O. Box 6025
West Virginia University
Morgantown, WV 26506-6025
Phone: (304) 293-7868
Email: dmitchel@wvu.edu

ABSTRACT

In linear-quadratic control (LQC) problems with singular control cost matrix and/or singular transition matrix, we derive a reduction of the dimension of the Riccati matrix, simplifying iteration and solution. Employing a novel transformation, we show that, under a certain rank condition, the matrix of optimal feedback coefficients is linear in the reduced Riccati matrix. For a substantive class of problems, our technique permits scalar iteration, leading to simple analytical solution. By duality the technique can also be applied to Kalman filtering problems with a singular measurement error covariance matrix.

JEL classification: C61; C63, D83

Keywords: Linear-quadratic control; Riccati equation; Riccati reduction; Kalman filtering; Intertemporal optimization

* The authors thank Victor Claar for stimulating their interest in the subject material.

Reducing the Dimensionality of Linear Quadratic Control Problems

1. Introduction

The preeminence of computable general equilibrium models has stimulated interest in the solution procedures for larger-scale models. Most commonly, linear rational expectations models are considered which, typically, are derivable from linear-quadratic control (LQC) problems. The recent work by Sims (2000), Binder and Pesaran (2000), King and Watson (1997, 1998), Amman (1997), Amman and Neudecker (1997), Anderson et al. (1997), Anderson and Moore (1985), Ehlgen (1999), and Klein (2000) concentrates on numerical procedures that (1) allow speedy and convenient computation of results, and (2) apply as generally as possible, in particular to systems with non-invertibilities stemming from a singular transition matrix or a singular control cost matrix. Generally, objective (2) has been attained at the expense of increasing the dimensionality of the system. These papers improve on the work of Vaughan (1970) and Blanchard and Kahn (1980).

The focus in the current paper is on simplifying the system when singularities occur. These simplifications lead to very simple analytical solutions in a substantial class of problems and in computationally efficient methods in all other cases, for finite or infinite horizon. We obtain a simple rank expression that places an upper bound on the effective dimensionality of the system for analytical and computational purposes: Prior computation of the rank of a composite matrix constructed from all coefficient matrices in the problem statement allows the researcher to establish this bound. The advantage is that one may readily determine up front whether the system has a simple explicit solution, or to what extent reformulation of the problem along the lines delineated here may reduce computation time or improve the transparency of the model.

The approach developed in this paper is aided by the duality between linear-quadratic

control and Kalman filtering as first described by Kalman (1960). A key feature of our approach is the use of the covariance matrices of Kalman filtering and their inverses (the latter being related to the Bayesian concept of precision). The Kalman-updated covariance matrix plays an important role that is not typically exploited in LQC analyses.

The Kalman approach also provides intuition for why the dimensionality of a particular system may be reduced. Consider the following class of Kalman problems, which is exemplified by Claar's (2000) model of the cyclical and natural unemployment rates. An observation depends linearly on two unobserved state variables following stochastic processes: $w_t = y_{1t} + y_{2t}$. One may describe the uncertainty of the state by considering the conditional variances σ_{1t}^2 , σ_{2t}^2 of the state variables and their conditional covariance $\sigma_{12t} = \sigma_{21t}$ (three numbers, stored in a 2×2 covariance matrix). However, conditional on having observed w_t , it is easy to derive that $\sigma_{1t}^2 = \sigma_{2t}^2 = -\sigma_{12t}$ so that one number is sufficient to describe the state uncertainty. This simplification was employed by Balvers and Cosimano (1994) in reducing the dimensionality of their active learning model, but the approach has not been systematically investigated. While the intuition for simplification here is straightforward, our rank expression implies a potentially complex interaction between the different singularities in the system that is not always intuitive.

Mitchell (2000) derived explicit analytical solutions to the 2×1 linear-quadratic control problem (two target variables and one uncosted control in the control case, or two state variables and one perfect observation in the Kalman case), but his results were not obviously generalizable. In this paper we significantly extend the class of linear-quadratic models which can be simplified or solved analytically in simple fashion.

The paper is organized as follows. Section 2 contains the theoretical approach in which

we derive theorems that state how the dimensionality of the model can be reduced and by how much, how under a certain rank condition the optimal feedback control matrix is linearly related to the reduced Riccati matrix (allowing a simple time-invariant linear transformation to obtain the feedback matrix once the appropriate reduced Riccati matrix is found), and how the reduced Riccati iterations can be replaced by linear iteration of a related matrix. Section 3 discusses implications for applied optimal control and Kalman filtering and presents some control and Kalman examples. In Section 4 we conclude the paper by providing a “how to” summary of our technique for practical use.

2. THEORY

2.1 *The control problem*

In this section we show how to reduce the dimension of the Riccati equation of optimal control, without assuming the presence of an invertible control cost matrix or invertibility of the transition matrix. In so doing we illuminate the underlying structure of the dynamics. The two initial lemmas establish the structure of the Riccati matrices, and Theorem 1 gives the reduced dynamics. The reduced problem is shown in Theorem 2 to be sometimes amenable to further simplification of the solution for the control feedback matrix. Then Theorem 3 shows how to obtain the Riccati sequence by linear iteration of a related matrix, permitting speedy calculation of the control sequence or its steady state. Theorems 1* and 3* deal with a further reduction of the Riccati matrix dimension which is possible under some conditions.

The reduction that we present is unrelated to the concept of reducing a system to “minimal” form in the sense of obtaining the lowest state vector dimension necessary for optimal Kalman filtering. A system is in minimal form if it is both controllable and observable (see Hannan and Deistler, 1988; note that some authors, especially in the Kalman filtering context, use

terminology different from ours, which is typically used in the control literature). In our control context, we assume that the state cost matrix is nonsingular, so that minimality is equivalent to controllability. Our Riccati reduction applies even if the system is controllable and hence minimal. This reduction can be achieved because the effective dimension of the Riccati matrix is less than the dimension of the state vector even if the latter is minimal.

A typical statement of the optimal control problem is:

$$(1a) \quad V(y_s, s) = \underset{\{u_t\}_{s+1}^T}{\text{Min}} \left(\frac{1}{2}(y_T' K_T y_T) + \sum_{t=s+1}^{T-1} [\frac{1}{2}(y_t' K y_t)] \right),$$

$$(1b) \quad \text{subject to } y_t = A y_{t-1} + C u_t, \quad t = s+1, \dots, T, \quad y_s \text{ given},$$

where the state cost matrices K and K_T , and the transition matrix A are $n \times n$, the control multiplier matrix C is $n \times k$, the state vector y_t is $n \times 1$, and the control vector u_t is $k \times 1$. The cost matrix K and the terminal cost matrix K_T are positive definite, and C has full column rank; A need not have full rank. If the original problem statement has control costs, one can augment the state vector with the costed controls [see Chow (1975)], putting all costs on the state vector and thus giving the problem formulation in equations (1).¹ It is well known [Chow (1975)] that the optimal controls are given by:

¹ In this case the reduction in this paper constitutes a net reduction of the Riccati matrix dimension if the original control cost matrix has less than full rank (or, even if not, often if the transition matrix has less than full rank).

Discounting and additive uncertainty in the state equation can easily be incorporated by redefining the transition matrix and the state and control vectors, and appealing to certainty equivalence, respectively. In addition, if the original problem statement has cross-product costs between the state vector and the control vector, these can be removed by redefinition of the state cost matrix and the control vector [see for instance Ljungqvist and Sargent (2000), p. 640].

$$(2) \quad u_t^{opt} = -(C' H_t C)^{-1} C' H_t A y_{t-1} \equiv -F_t y_{t-1}, \quad t \leq T,$$

$$(3) \quad H_{t-1} = K + A' H_t A - A' H_t C (C' H_t C)^{-1} C' H_t A, \quad H_T = K_T, \quad t \leq T,$$

where the symmetric $n \times n$ matrix H_t is positive definite.

2.2 The basic reduction

Equation (3) can be written as equations (4) and (5):

$$(4) \quad H_{t-1} = K + A' P_t A, \quad t \leq T,$$

$$(5) \quad P_t = H_t - H_t C (C' H_t C)^{-1} C' H_t, \quad t \leq T.$$

The symmetric $n \times n$ P_t matrix is not typically employed in dealing with optimal control problems, but in the dual Kalman filtering context has the familiar interpretation of the covariance matrix for the unobserved state variable for the current period conditional on current information (while H_t is the covariance matrix conditional on the previous period's information).

The approach in this paper is to exploit restrictions inherent in the P_t matrix to simplify the solution of problem (1). By equation (5) we have:

$$(6) \quad P_t C \equiv \begin{bmatrix} P_{1t} & P_{2t} \\ P_{3t} & P_{4t} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = 0_{n,k}, \quad t \leq T,$$

where, defining $q = n - k$, we have that P_{1t} is $q \times q$, P_{2t} is $q \times k$, $P_{3t} = P_{2t}'$, and P_{4t} is $k \times k$; C_1 is $q \times k$, and C_2 is $k \times k$. P_{1t} and P_{4t} are symmetric.

Since C is of full column rank, there is at least one $k \times k$ sub-matrix of C that is invertible. Proper prior arrangement of the y_t vector (and concomitant arrangement of C , A , K , and K_T) is thus sufficient to guarantee that C_2 is invertible. We can then derive:

LEMMA 1 (REDUCTION TO THE DYNAMIC CORE OF P_t). The $n \times n$ matrix

P_t can be written as:

$$(7) \quad P_t = M \Phi_t M' , \quad \Phi_t \equiv P_{1t} , \quad t \leq T$$

$$(8) \quad M \equiv \begin{bmatrix} I_{n-k} \\ -(C_2')^{-1} C_1' \end{bmatrix} ,$$

where M is an $n \times q$ matrix, and Φ_t is invertible with dimensions $q \times q$.

Proof. From equations (6) it is straightforward to relate P_{3t} ($= P_{2t}'$) and P_{4t} to P_{1t} . The first equation in (6) gives $P_{2t} = -P_{1t} C_1 C_2^{-1}$. Transpose (to produce P_{3t}) and substitute into the second equation (noting the symmetry of P_{1t} as follows from the symmetry of P_t) which yields $P_{4t} = (C_2')^{-1} C_1' P_{1t} C_1 C_2^{-1}$. Then factor out the M and M' matrices to produce equation (7).

To show that Φ_t is invertible, note from equation (5) that P_t can be written as the product

$H_t [I_n - C(C'H_t C)^{-1} C'H_t]$, where the matrix in brackets is idempotent with trace equal to $trace(I_n) - trace(I_k)$ and thus rank $n - k = q$. Thus, since H_t has full rank n , P_t has rank q .² Equation (7) then implies that $\text{rank}(\Phi_t) \geq q$, and since Φ_t has dimension q it must have full rank.

■

It will be important to relate Φ_t to the solution of problem (1) – equations (2) and (3) – in a meaningful way. Lemma 2 provides a useful link.

LEMMA 2 (RELATING Φ_t AND H_t). The $q \times q$ matrix Φ_t in equation (7) is positive definite and is given by:

$$(9) \quad \Phi_t = (M'H_t^{-1}M)^{-1}, \quad t \leq T.$$

Proof. Post-multiply equation (5) by $H_t^{-1}P_t$ and then use the transpose of equation (6). This yields $P_t = P_t H_t^{-1} P_t$, so that, interestingly, H_t^{-1} is seen to be a generalized inverse of P_t . Next use equation (7) in the right-hand side of this equation and pre-multiply by $\begin{bmatrix} I_q & 0 \end{bmatrix}$ and post-multiply by $\begin{bmatrix} I_q & 0 \end{bmatrix}'$ to pick out the upper left block $P_{1t} \equiv \Phi_t$ of the matrix, yielding:

$$(10) \quad \Phi_t = \begin{bmatrix} I_q & 0 \end{bmatrix} M \Phi_t M' H_t^{-1} M \Phi_t M' \begin{bmatrix} I_q \\ 0 \end{bmatrix}, \quad t \leq T.$$

² P_t has rank q by Sylvester's inequality: $\text{rank}(X_1) + \text{rank}(X_2) - n \leq \text{rank}(X_1 X_2) \leq \min[\text{rank}(X_1), \text{rank}(X_2)]$, where n is the number of rows in X_2 .

Now consider that $\begin{bmatrix} I_q & 0 \end{bmatrix} M = I_q$, and post-multiply equation (10) by $\Phi_t^{-1} (M' H_t^{-1} M)^{-1}$, to obtain equation (9). Positive definiteness follows directly from equation (9) given that H_t is positive definite. ■

Employing Lemmas 1 and 2 we now provide the dynamics of Φ_t .

THEOREM 1 (DYNAMICS OF Φ_t^{-1}). For all $t \in \{s+1, T\}$ we have :

$$(11) \quad \Phi_{t-1}^{-1} = B_1 - B_2' (\Phi_t^{-1} + B_3)^{-1} B_2, \quad \Phi_T^{-1} = M' K_T^{-1} M,$$

$$\text{with } B_1 = M' K^{-1} M, \quad B_2 = M' A K^{-1} M, \quad B_3 = M' A K^{-1} A' M.$$

Proof. Substitute $P_t = M \Phi_t M'$ from Lemma 1 into equation (4):

$$(12) \quad H_{t-1} = K + A' M \Phi_t M' A, \quad t \leq T.$$

A standard inversion identity (used later on further occasions) gives:³

$$(13) \quad H_{t-1}^{-1} = K^{-1} - K^{-1} A' M (\Phi_t^{-1} + M' A K^{-1} A' M)^{-1} M' A K^{-1}.$$

³ Given the matrices X_1, X_2, X_3 , and X_4 , with X_1 and X_4 invertible, we have the identity [Söderström (1994), pp. 156-7]:

$$(X_1 + X_2 X_4^{-1} X_3)^{-1} = X_1^{-1} - X_1^{-1} X_2 (X_4 + X_3 X_1^{-1} X_2)^{-1} X_3 X_1^{-1}.$$

Post-multiplying by M and pre-multiplying by M' yields after applying Lemma 2:

$$(14) \quad \Phi_{t-1}^{-1} = M' K^{-1} M - M' K^{-1} A' M (\Phi_t^{-1} + M' A K^{-1} A' M)^{-1} M' A K^{-1} M,$$

which is equation (11). $\Phi_T^{-1} = M' K_T^{-1} M$ follows from equation (9) using $H_T = K_T$ from (3). ■

By employing the matrix inversion identity of footnote 3 to equation (11), we obtain the following dynamics for Φ_t :

COROLLARY 1.1 (DYNAMICS OF Φ_t). For all $t \in \{s+1, T\}$ we have

$$(15) \quad \Phi_{t-1} = B_1^{-1} + B_1^{-1} B_2' [\Phi_t^{-1} + (B_3 - B_2 B_1^{-1} B_2')]^{-1} B_2 B_1^{-1}.$$

2.3 The case of nonsingular B_2

Note that the B_i matrices in Theorem 1 are all $q \times q$ and that only B_2 is not symmetric. B_1 is positive definite and B_3 is positive semi-definite. By Sylvester's inequality (see footnote 2), B_2 ($= M' A K^{-1} M$) can be of full or less than full rank regardless of whether A has full rank.

(However, if $\text{rank}(A) < q$ then B_2 is certainly singular.) If B_2 is invertible we obtain additional results; if B_2 is singular, further reduction is shown to be possible and results analogous to some of those for the case of nonsingular B_2 are obtained.

COROLLARY 1.2 (REVERSIBILITY). If B_2 is invertible, the time direction

of the recursion for $\{\Phi_t^{-1}\}$ in equation (11) can be reversed:

$$(16) \quad \Phi_t^{-1} = B_2 (B_1 - \Phi_{t-1}^{-1})^{-1} B_2' - B_3, \quad t \leq T.$$

Thus, using equation (7) in equation (4) to get H_{t-1} in terms of Φ_t , the recursion for $\{H_t\}$ is reversible as well, implying that there is a one-to-one correspondence between H_t and H_T for any t . In contrast the reversibility is not at all clear in the standard Riccati equation (3).

The sequence of reduced Riccati matrices obtained in Theorem 1 can be used with equations (2), (4), and (7) to obtain the sequence $\{u_t^{opt}\}$ of optimal controls. However, given the transformations employed here there is a more convenient way of calculating the optimal controls when B_2 has full rank ($=q$), as in this case the feedback matrix can be shown to be linear in Φ_t :

THEOREM 2 (LINEAR CALCULATION OF FEEDBACK MATRIX). Given

$u_t^{opt} = -F_t y_{t-1}$, if $rank(B_2) = q$, then the feedback matrix F_t is linear in

Φ_t for all $t \leq T - 1$:

$$(17) \quad F_t = -WM\Phi_t M' A + WKA, \quad t \leq T - 1,$$

$$\text{with } W = (C'C)^{-1} C' K^{-1} [I - A'M(M'K^{-1}A'M)^{-1} M'K^{-1}].$$

Proof. From equations (2) and (5) we obtain

$$(18) \quad CF_t = (I_n - H_t^{-1} P_t) A = (I_n - H_t^{-1} M \Phi_t M') A, \quad t \leq T,$$

where the second equality follows from Lemma 1. To obtain the term $H_t^{-1}M$ appearing on the right-hand side of equation (18), we first use equation (13) and the definitions in Theorem 1:

$$(19) \quad H_{t-1}^{-1}M = K^{-1}M - K^{-1}A'M(\Phi_t^{-1} + B_3)^{-1}B_2, \quad t \leq T.$$

Use the solution of equation (11) for $(\Phi_t^{-1} + B_3)^{-1}$ in equation (19):

$$(20) \quad H_{t-1}^{-1}M = K^{-1}M - (K^{-1}A'M)B_2^{-1}'(B_1 - \Phi_{t-1}^{-1}), \quad t \leq T.$$

Update equation (20) by one period (making it valid for $t \leq T - 1$) and substitute into the right side of (18). Pre-multiplying the left and right sides of equation (18) by $(C'C)^{-1}C'$ yields equation (17). ■

Thus computation of the sequence $\{F_t\}$ of control feedback matrices involves first computing $F_T = (C'K_T C)^{-1}C'K_T A$ from equation (2) with $H_T = K_T$, next iterating equations (11) or (15) to get $\{\Phi_t\}$, and then using equation (17) to obtain the remainder of the feedback matrix sequence.

Comparing to the solution of Problem (1) in equations (2) and (3), the solution proposed here in Theorem 1 [equation (11)] and Theorem 2 [equation (17)] is computationally advantageous. The size of the Φ_t matrix is $q = n - k$, whereas the size of H_t is n . And the iterations in this approach require fewer matrix multiplications – two in equation (11) and two in equation (17) (aside from the fixed cost of initial matrix calculations) – while the traditional

approach requires seven multiplications in equation (3) and no additional multiplications in equation (2).

The nonlinear recursive equation (11) for Φ_t can be replaced by a linear equation of larger dimension by using a power method [see for instance Martin and Ammar (1991)]:

THEOREM 3 (LINEAR ITERATION). For any $t \in \{s, T\}$ we have for the case of $\text{rank}(B_2) = q$:

$$(21) \quad \Phi_t = N_t D_t^{-1},$$

with

$$(22) \quad \begin{bmatrix} N_{t-1} \\ D_{t-1} \end{bmatrix} = B \begin{bmatrix} N_t \\ D_t \end{bmatrix}, \quad N_T = I_q, \quad D_T = M' K_T^{-1} M,$$

and hence

$$(23) \quad \begin{bmatrix} N_t \\ D_t \end{bmatrix} = (B)^{T-t} \begin{bmatrix} I_q \\ M' K_T^{-1} M \end{bmatrix},$$

where

$$(24) \quad B = \begin{bmatrix} B_2^{-1} B_3 & B_2^{-1} \\ B_1 B_2^{-1} B_3 - B_2' & B_1 B_2^{-1} \end{bmatrix}.$$

Here the B_i are defined in Theorem 1, and N_t and D_t are $q \times q$.

Proof. For B_2 invertible, we can post-multiply both sides of equation (11) by the inverse of $(\Phi_t^{-1} + B_3)^{-1} B_2$ and then pre-multiply through by Φ_{t-1} and post-multiply through by Φ_t to produce:

$$(25) \quad B_2^{-1} - \Phi_{t-1} B_1 B_2^{-1} + B_2^{-1} B_3 \Phi_t + \Phi_{t-1} (B_2' - B_1 B_2^{-1} B_3) \Phi_t = 0, \quad t \leq T.$$

[Interestingly, the steady-state version of equation (25) (with time subscripts omitted) is of the same form as the algebraic Riccati equation of linear-quadratic control in *continuous* time (Lancaster and Rodman, 1995)]. Solve equation (25) for Φ_{t-1} :

$$(26) \quad \Phi_{t-1} = [B_2^{-1} + B_2^{-1} B_3 \Phi_t] [B_1 B_2^{-1} + (B_1 B_2^{-1} B_3 - B_2') \Phi_t]^{-1}, \quad t \leq T.$$

Postulate $\Phi_t = N_t D_t^{-1}$ (which is true for $t = T$ with $N_T = \Gamma$, $D_T = (M' K_T^{-1} M) \Gamma$ for arbitrary invertible matrix Γ , which we set to I_q), use this in equation (26), and post-multiply all four terms by D_t to get

$$\Phi_{t-1} = [B_2^{-1} D_t + B_2^{-1} B_3 N_t] [B_1 B_2^{-1} D_t + (B_1 B_2^{-1} B_3 - B_2') N_t]^{-1}, \quad t \leq T.$$

Then N_{t-1} and D_{t-1} are the first and second bracketed terms here, confirming equation (22) by induction. Equation (23) follows directly from equation (22). ■

To calculate the steady state or “infinite horizon” value, $\Phi_{-\infty} \equiv \lim_{t \rightarrow -\infty} N_t D_t^{-1}$,

one may just iterate equation (22). Note that both N_t and D_t will generally explode but that

$N_t D_t^{-1}$ remains finite. To prevent computational problems it may be advisable to rescale

periodically in finding $\Phi_{-\infty}$ by multiplying both N_t and D_t by a common scalar, leaving $N_t D_t^{-1}$

unaffected. The iterations to the steady state can be speeded by using a doubling algorithm, as in

Anderson and Moore (1979, pp. 159-160). We successively compute $B^2, (B^2)^2, \dots, B^{2^m}$,

rescaling as needed. Then $a \begin{bmatrix} N_{T-2^m} \\ D_{T-2^m} \end{bmatrix} = a B^{2^m} \begin{bmatrix} N_T \\ D_T \end{bmatrix}$, where a is the rescaling factor.

2.4 Further reduction when B_2 is singular

We now consider the case in which B_2 in equation (11) is singular. This provides the opportunity for further reduction of the size of the Riccati matrix: the $n \times n$ Riccati matrix H has already been reduced to the $q \times q$ Riccati matrix Φ ; if the $q \times q$ matrix B_2 has rank $r < q$, we can further reduce Φ to an $r \times r$ matrix to be denoted Φ^* .

First put the $q \times q$ matrix B_2 in standard form:

$$I_r^q = Q B_2 S, \quad \text{where } I_r^q \equiv \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix}.$$

Here Q and S are invertible $q \times q$ matrices and S_i is $r \times r$. S must be arranged such that S_4 is invertible (which may require a row and column rearrangement as discussed in footnote 4).

Then we have:

THEOREM 1* (FURTHER REDUCTION OF Φ_t). If B_2 has less than full rank,

the dimension of the time-varying core of the Riccati equation, denoted

Φ_t^* , is $r \times r$, where $r = \text{rank}(B_2)$. Φ_t^* is positive definite, with

$\Phi_t^{*-1} \equiv Z' \Phi_t^{-1} Z$ where Z is defined below, and its dynamics is described

by:

$$(27) \quad \Phi_{t-1}^{*-1} = B_1^* - B_2^{*'} (\Phi_t^{*-1} + B_3^*)^{-1} B_2^* , \quad t \leq T-1,$$

$$\Phi_{T-1}^{*-1} = Z' [B_1 - B_2' (\Phi_T^{-1} + B_3)^{-1} B_2] Z;$$

$$\text{with } B_1^* = Z' \{ B_1 - B_2' (B_1 + B_3)^{-1} B_2 + B_2' (B_1 + B_3)^{-1} \cdot$$

$$M^* [M^{*'} (B_1 + B_3)^{-1} M^*]^{-1} M^{*'} (B_1 + B_3)^{-1} B_2 \} Z,$$

$$B_2^* = [M^{*'} (B_1 + B_3)^{-1} M^*]^{-1} M^{*'} (B_1 + B_3)^{-1} B_2 Z ,$$

$$B_3^* = [M^{*'} (B_1 + B_3)^{-1} M^*]^{-1} - Z' B_1 Z,$$

$$\text{where } Z = \begin{bmatrix} I_r \\ 0_{q-r, r} \end{bmatrix} \text{ and } M^* = \begin{bmatrix} I_r \\ -S_4^{-1'} S_2' \end{bmatrix} \quad (\text{so } M^* \text{ is } q \times r).$$

Proof. In Appendix A .

In equation (27), B_2^* may or may not be invertible. (Appendix A gives an example in

which it is not.) In either case, Theorem 1* may be combined with equations (2), (4), (7), and

(A3) to obtain the sequence of optimal controls based on $\{\Phi_t^*\}$. If B_2^* is not invertible, we can further reduce the dimensionality of the problem by repeatedly applying the reduction process of Theorem 1* until the reduced analog of B_2^* either has zero rank or is invertible. Thus the effective dimension of the original problem is less than or equal to the rank of B_2 .

Note that the proof of Theorem 1* required that the initial condition be stated as of period $T - 1$. Thus in iterating H_t , H_T equals K_T , while for $t \leq T - 1$ Φ_t^* is used to compute Φ_t , which in turn is used to compute H_t . Any further reduction shifts back one more period the time of the initial condition for reduced Riccati iteration.

There is no counterpart to Theorem 2 for the case in which any reduction beyond that of Theorem 1 has been applied, because the proof of Theorem 2 rests crucially on invertibility of B_2 and cannot be generalized to rely on invertibility of B_2^* . In this case of singular B_2 we can, however, obtain linear iteration analogous to that of Theorem 3 for the nonsingular case:

THEOREM 3* (LINEAR ITERATION OF Φ_t^*). For any $t \in \{s, T\}$ we have for the case of $\text{rank}(B_2) < q$, when Theorem 1* has been applied until B_2^* is invertible:

$$\Phi_t^* = N_t^* D_t^{*-1},$$

with

$$(28) \quad \begin{bmatrix} N_{t-1}^* \\ D_{t-1}^* \end{bmatrix} = B^* \begin{bmatrix} N_t^* \\ D_t^* \end{bmatrix}, \quad N_T^* = I_r, \quad D_T^* = Z' M' K_T^{-1} M Z,$$

and hence

$$(29) \quad \begin{bmatrix} N_t^* \\ D_t^* \end{bmatrix} = (B^*)^{T-t} \begin{bmatrix} I_r \\ Z' M' K_T^{-1} M Z \end{bmatrix},$$

where

$$(30) \quad B^* = \begin{bmatrix} B_2^{*-1} B_3^* & B_2^{*-1} \\ B_1^* B_2^{*-1} B_3^* - B_2^{*'} & B_1^* B_2^{*-1} \end{bmatrix}.$$

The B_i^* are defined in Theorem 1*.

Proof. In Appendix B.

This section has shown how to reduce the size of the dynamic Riccati matrices of optimal control, thereby simplifying computation and revealing the underlying structure of the dynamics. To obtain these results we did not require invertibility either of the matrix of control costs or of the transition matrix A . In addition we have provided simple linear dynamics. The usefulness of these results will be demonstrated in the applications section to follow.

3. Implications and Applications

3.1 Effective dimension of the system

By Theorem 1*, the upper bound on the effective dimension of the system (the size of Φ_t^* or of Φ_t if B_2 has full rank) is given by the rank of $B_2 \equiv M'AK^{-1}M$ with M defined in equation (8). This bound may be determined in advance – that is, before theoretical appraisal, estimation, or numerical analysis, or explicit solution of the model. A general indication of the rank of B_2 is obtained by repeated application of Sylvester's inequality (presented in footnote 2) to the definition of B_2 given in Theorem 1. Recalling that n represents the dimension of the state vector and k the number of controls, Sylvester's inequality yields:

$$(31) \quad \text{rank}(A) - 2k \leq \text{rank}(B_2) \leq \min[n - k, \text{rank}(A)].$$

Scalar Riccati dynamics will be guaranteed if $n - k$ (the size of B_2) = 1 (or, of course, if $\text{rank}(A) = 1$). This case will be discussed in the next sub-section.

Before we discuss the scalar case, we present a simple example to illustrate the bounds implied by equation (31). Consider a case with $n = 3$, $k = 1$, $C' = [0 \ 0 \ 1]$, and $K = I_3$. The 3×3 matrix A is unrestricted. (Note that it is always possible, starting from any like-sized problem with any C , to transform the control and state vectors so that $C' = [0 \ 0 \ 1]$.) Then, $B_2 = A_1$, where A_1 is the 2×2 upper left block of A . Consequently, there is an infinitude of A matrices for which: (a) $\text{rank}(B_2) = \text{rank}(A_1) = n - k = 2$, when the two 1×2 rows of A_1 are independent, with $\text{rank}(A)$ equaling either three or two; (b) $\text{rank}(B_2) = \text{rank}(A_1) =$

$\text{rank}(A) - 2k = 1$, when A has rank one, two, or three and the two 1×2 rows of A_1 are dependent; and (c) $\text{rank}(B_2) = \text{rank}(A_1) = \text{rank}(A) - 2k = 0$, when all four elements of A_1 are zero so that A must be singular with either rank one or two.

3.2 Analytical solution when $\text{rank}(B_2) \leq 1$

When $B_2 = M' A K^{-1} M$ has rank equal to or less than one, the LQC problem allows scalar-based analytical solution. When the rank of B_2 is equal to *zero* because $B_2 = 0$, Theorem 1 directly shows that Φ_t does not evolve. When the rank of B_2 is equal to *one*, Theorem 1 applies if $n - k = 1$ (so B_2 has full rank) and Theorem 1* applies if $n - k > 1$ (so B_2 has less than full rank). In what follows we discuss the case $n - k = 1$, but if the rank of B_2 is less than full and equals 1, Theorem 1* applies and the results below all continue to hold if we replace the B_i by B_i^* . Equation (11) implies the scalar equation:

$$(32) \quad \Phi_{t-1} = [1 + B_3 \Phi_t] / [B_1 + (B_1 B_3 - B_2^2) \Phi_t], \quad t \leq T,$$

where Theorem 1 defines B_1 , B_2 , and B_3 , which are scalar in this case (and if Theorem 1* applies then B_1^* , B_2^* , and B_3^* are scalar).

Mitchell (2000) shows how to find the solution to a scalar equation of the form of equation (32) as follows. Consider first the case of $B_1 B_3 - B_2^2 \neq 0$, so that Φ_t evolves nonlinearly (unless $B_2 = 0$). Let $x_t = 1/(c + \Phi_t)$ and hence $\Phi_t = (1 - cx_t)/x_t$, where

$c = (B_1 - B_3 + r)/[2(B_1B_3 - B_2^2)]$ and $r = [(B_1 - B_3)^2 + 4(B_1B_3 - B_2^2)]^{1/2}$. Then use

$\Phi_t = (1 - cx_t)/x_t$ on both sides of equation (32) to obtain a linear equation of evolution for x_t :

$$(33) \quad x_{t-1} = \frac{2(B_1B_3 - B_2^2)}{B_1 + B_3 + r} + \left(\frac{B_1 + B_3 - r}{B_1 + B_3 + r} \right) x_t, \quad t \leq T,$$

with solution

$$(34) \quad x_t = \frac{B_1B_3 - B_2^2}{r} + \left(x_T - \frac{B_1B_3 - B_2^2}{r} \right) \left(\frac{B_1 + B_3 - r}{B_1 + B_3 + r} \right)^{T-t}, \quad t \leq T.$$

Then the solution for Φ_t is found by putting equation (34) into $\Phi_t = (1 - cx_t)/x_t$.

It is also possible for equation (32) to give *linear* evolution of Φ_t . This occurs if and only if $B_1B_3 - B_2^2 = 0$. In this linear case the solution of equation (32) for Φ_t is obvious and the eigenvalue is B_3/B_1 , which [as Mitchell (2000) shows] may or may not be less than one in magnitude so the linear case may or may not be stabilizable.

To examine the nature of the scalar dynamics, first derive from equation (32):

$$(35) \quad d\Phi_{t-1}/d\Phi_t = B_2^2/[B_1 + (B_1B_3 - B_2^2)\Phi_t]^2 \geq 0.$$

Equation (35) suggests several different cases:

Case 1: $B_2 = 0$. This case is covered equally well by Theorem 1 or Theorem 1*. Equation (32) collapses to $\Phi_{t-1} = 1/B_1$ which is constant and equal to the endpoint value Φ_T if $K_T = K$. Figure 1(a) shows the dynamics of Φ_t and hence H_t if $K_T \neq K$ so $\Phi_T \neq 1/B_1$: the steady state is reached in one iteration.

Case 2: $B_2 \neq 0$ and $B_1 B_3 - B_2^2 \neq 0$. Note that $B_1 B_3 - B_2^2$ cannot be negative: we know $B_3 - B_2 B_1^{-1} B_2' = M' A K^{-1} [K - M(M' K^{-1} M)^{-1} M'] K^{-1} A' M = M' A C (C' K C)^{-1} C' A' M$, where the last equality follows from substituting equation (9) into equation (7) and the result into equation (5), evaluating the resulting identity at $H_t = K$, subtracting K from both sides, and pre- and post-multiplying both sides by K^{-1} . Hence, $B_3 - B_2 B_1^{-1} B_2'$ is positive semi-definite, and so in this scalar case multiplying this expression by the positive scalar B_1 establishes $B_1 B_3 - B_2^2 \geq 0$. Then in this case 2 equation (35) implies that $d\Phi_{t-1}/d\Phi_t > 0$ and $d^2\Phi_{t-1}/d\Phi_t^2 < 0$; and as $\Phi_t \rightarrow \infty$ we have $d\Phi_{t-1}/d\Phi_t \rightarrow 0$. Thus, the time path is monotonic and convergent as displayed in Figure 1(b).

Case 3: $B_2 \neq 0$ and $B_1 B_3 - B_2^2 = 0$. Now by equation (32), $\Phi_{t-1} = (1/B_1) + (B_3/B_1) \Phi_t$ so evolution is linear. This permits the stable case of $B_3 < B_1$ shown in Figure 1(c) (noting that both $B_1 (= M' K^{-1} M)$, and $B_3 (= M' A K^{-1} A' M)$ must be nonnegative given positive definite K) as well as the unstable case of $B_3 \geq B_1$, also shown in Figure 1(c).

We have shown here how to solve the case of $n - k = 1$ analytically, which was heretofore

done only for the $n = 2, k = 1$ case by Mitchell (2000). In addition, we have shown how, due to potential singularities in the transition matrix and its interactions with the cost matrix, other apparently more complex problems can also be solved analytically if the effective dimensionality is 1.

3.3 The Amman and Neudecker example

An example of Amman and Neudecker (1997) can be translated into the framework of equations (1) by moving their control cost matrix into an augmented state cost matrix, augmenting the controls into the state vector, augmenting the transition matrix A with a row and column of zeroes, and augmenting the C matrix with a one to equate identically the new state variable to the control variable. Thus

$$A = \begin{bmatrix} 0.8 & 0 & 0 & 0 & -0.8 & 0 \\ 0.8 & 0 & 0 & 0 & -0.8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad K = I_6, \quad C = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

In this case, the B_2 matrix can be seen to equal the principal 5×5 minor of the A matrix, which is the original transition matrix in Amman and Neudecker. The rank of B_2 is equal to 2; since B_2^* turns out to have full rank, the effective dimensionality of the problem is 2.

Theorem 1* can be applied to produce the sequence $\{\Phi_t^*\}$ and hence $\Phi_{-\infty}^*$ and $\Phi_{-\infty}$.

Equations (4) and (7) then give $H_{-\infty}$, of which the upper left 5×5 block is identical to the steady

state Riccati matrix in Amman and Neudecker (1997). While both our method and that of Amman and Neudecker converge very quickly, ours has the advantage of simultaneously yielding the non-steady-state iterates of the system.

3.4 Kalman filtering

The duality between LQC problems and Kalman filtering was first noted by Kalman (1960). Accordingly, our theoretical results apply in the Kalman filtering context as well.

Consider the following class of state space models:

$$(36a) \quad y_{\tau-1} = A' y_{\tau} + \eta_{\tau}, \quad \eta_{\tau} \sim N(0, K),$$

$$(36b) \quad w_{\tau} = C' y_{\tau}, \quad \tau \leq T,$$

where w_{τ} represents a $k \times 1$ vector of observations and y_{τ} is an $n \times 1$ vector of unobservable state variables. The interpretation of τ here is “periods left until the last observation.” The disturbances in the state equations are serially uncorrelated and multivariate normal, with covariance matrix K . The dimensions of the coefficient matrices are as in the LQC problem.

Note that the omission of measurement error in equation (36b) is analogous to the omission of control costs in the LQC model. Measurement error in some of the observations may be incorporated by extending the state vector. The possibility for reducing the dimensionality of the Riccati matrix is quite intuitive here and generalizes the approach of Balvers and Cosimano (1994) in a related case with two state variables and one observation. In their model

$w_{1\tau} = y_{1\tau} + c y_{2\tau}$. While the Riccati matrix has dimension 2×2 , containing the variances of $y_{1\tau}$ and $y_{2\tau}$ on the diagonal and their covariance off the diagonal, the single scalar $\Phi_{\tau} \equiv P_{1\tau}$

contains all variance-covariance information since, given that the observation is known with certainty, the observation equation implies: $P_{1\tau} = c^2 P_{4\tau} = -c P_{2\tau} = -c P_{3\tau}$. This is the intuition for equation (6).

The solution to the Kalman filtering problem modeled in equations (36) consists of a sequence of conditional expectations of the state variables $\{E_\tau y_{\tau-1}\}$ and their conditional covariance matrices $\{H_\tau\}$, for given priors in the initial period T . The iterative solution for $\{E_\tau y_{\tau-1}\}$ is [see for instance Harvey (1989)]:

$$(37) \quad E_\tau y_{\tau-1} = A' E_{\tau+1} y_\tau + F_\tau' \epsilon_\tau, \quad \epsilon_\tau \equiv w_\tau - C' E_{\tau+1} y_\tau, \quad \tau \leq T,$$

where F_τ is given in equation (17), which is linear in Φ_τ , with evolution as in equation (11). The sequence $\{H_\tau\}$ is given by equation (3), but may be found more easily from equation (11) with use of equations (4), (7), and (8).

Our reduction results therefore directly apply to problem (36). Accordingly, $rank(B_2)$ provides an upper bound on the relevant dimension of the problem for solution and (typically maximum likelihood) estimation purposes.

3.5 A specific Kalman filtering example when $rank(B_2) \leq 1$

The example below contains the models in Claar (2000) and Balvers and Wu (2000) as special cases. Claar considers the observed unemployment rate as consisting of the unobserved natural and cyclical unemployment rates. Balvers and Wu consider observed stock return indices in 18 countries as having a common unobserved world component and an unobserved

idiosyncratic component. Here, the Riccati matrix would be 19×19 , including all idiosyncratic components plus the common component. But there are 18 observations so that the dimension of B_2 is unity, and so $\text{rank}(B_2) \leq 1$.

Problem (36) is now specialized to the following coefficient matrices:

$$(38) \quad A = \begin{bmatrix} a_0 & 0 & \dots & 0 \\ 0 & a_1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{n-1} \end{bmatrix}, \quad K = \begin{bmatrix} k_0 & 0 & \dots & 0 \\ 0 & k_1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & k_{n-1} \end{bmatrix}, \quad C = \begin{bmatrix} 1 & \dots & 1 \\ \text{---} & \text{---} & \text{---} \\ 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}.$$

Note that both A and K are $n \times n$, whereas C is $n \times (n-1)$. Thus, for the Balvers and Wu example the first row of each of the coefficient matrices refers to the “common” component; there are $n-1$ idiosyncratic components, one for each observation. For the Claar example, $n = 2$ and so

$$C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

From equation (8) we can obtain the M matrix with dimension $n \times 1$:

$$(39) \quad M = [1 \quad -1 \quad -1 \quad \dots \quad -1]'$$

From equations (38) and (39) we straightforwardly calculate the scalar coefficients of the Φ_τ recursion of equation (11):

$$(40) \quad B_1 = \sum_{i=0}^{n-1} k_i^{-1}, \quad B_2 = \sum_{i=0}^{n-1} a_i k_i^{-1}, \quad B_3 = \sum_{i=0}^{n-1} a_i^2 k_i^{-1}.$$

The cases of Figures 1(a) - 1(c) all arise here. For brevity we state here only that Case 3 of Figure 1(c) can be ruled out except for the uninteresting case where all coefficients are equal. Case 1 of Figure 1(a) occurs if $B_2 = 0$. In this case, all there is to be learned is completed in one period. Otherwise, we have Case 2 in Figure 1(b).

To illustrate Theorem 2, we find the Kalman gain matrix F'_τ from equation (17). Simple but tedious calculation shows that there are three basically different forms of the Kalman gain coefficients $F'_{ji\tau} = F_{ij\tau}$:

$$(41) \quad F_{ij\tau} = \frac{\partial E_{\tau+1} x_{j-1\tau}}{\partial \epsilon_{i\tau+1}} = \begin{cases} -a_{j-1} z_{i\tau}, & j \neq 1, \quad j \neq i+1 \\ +a_0 z_{i\tau}, & j = 1 \\ -a_i (z_{i\tau} - 1), & j = i+1, \end{cases}$$

where
$$z_{i\tau} = \frac{a_i k_i^{-1}}{B_2} + \Phi_\tau k_i^{-1} \left(1 - \frac{a_i B_1}{B_2} \right).$$

It can be shown that, for proper priors, $z_{i\tau}$ is always positive in non-trivial cases. Thus, equation (41) implies that a surprise increase in observation $w_{i\tau+1}$ causes an increase in the expected value of next period's common component $y_{0\tau}$ and thus a decrease in the expected value of next period's idiosyncratic components $y_{j-1,\tau}$ unless $j = i+1$. In this case, $y_{i\tau}$,

representing the idiosyncratic component in the observation i itself, may decrease when uncertainty Φ_τ is high, but eventually increases as Φ_τ reaches its steady state. The intuition is that, with high uncertainty, the estimate of the common component $y_{0\tau}$ may be revised substantially upwards so that the estimate of the idiosyncratic component $y_{i\tau}$ could fall even though the observation $w_{i\tau} = y_{0\tau} + y_{i\tau}$ increases.

4. Summary and Conclusion

This paper presents a procedure for simplifying and solving LQC models. The procedure can be summarized in the following algorithm:

Step 1. If necessary, transform the LQC problem to fit the structure of equations (1).

Step 2. First obtain M from equation (8) and subsequently obtain B_1 , B_2 , B_3 , and Φ_T^{-1} as given in Theorem 1.

Step 3. If B_2 has full rank find $\{\Phi_i\}$ from Theorem 1 (or Theorem 3). Then skip to *Step 8*.

Step 4. If B_2 does not have full rank, find the S matrix by transforming B_2 into standard form and extract S_4^{-1} and S_2 . (If S_4^{-1} does not exist consider footnote 4.)

Step 5. Obtain M^* , B_1^* , B_2^* , B_3^* , and Φ_{T-1}^{*-1} as given in Theorem 1*. If B_2^* is not invertible repeat

Steps 4 and 5.

Step 6. Find $\{\Phi_i^{*-1}\}$ from Theorem 1* or 3*.

Step 7. Employ equation (A3) to deduce $\{\Phi_i\}$ from $\{\Phi_i^{*-1}\}$.

Step 8. Substitute $\{\Phi_t\}$ into equation (7) to find $\{P_t\}$ and then use equation (4) to generate

$\{H_t\}$, if $\{P_t\}$ and $\{H_t\}$ are needed. The end matrix H_T is generated as $H_T = K_T$, and

H_{T-1} is generated from Φ_T via equations (4) and (7).

Step 9. If B_2 has full rank, find the feedback matrix sequence $\{F_t\}$ from equation (2), or, for

$t \leq T - 1$, from Theorem 2. In the LQC problem, the optimal control vector u_t^{opt} equals

$-F_t y_{t-1}$. In the Kalman problem F_t' is the information filter appearing in equation (37).

Step 9'. If B_2 is singular, use $\{H_t\}$ from *Step 8* and use equation (2) to find the $\{F_t\}$ matrix sequence.

This procedure has a series of advantages compared to traditional procedures for solving the finite horizon and infinite horizon LQC problem, such as Vaughan (1970), and compared to more recent procedures. First, it provides a simple calculation (the rank of B_2) to establish an upper bound on the effective dimension of the problem. It is then possible to find in advance, without computing the solution, how complicated or simple the dynamics is or how much computing time may be saved by reformulating the problem. Second, the procedure leads to scalar-based analytical solutions when the rank of B_2 is zero or one, and to matrix size reductions in other cases when the transition matrix and the matrix of control costs may be less than full rank. Third, in the general $n \times k$ problem, iterations can be conducted with matrices that have the size of the effective dimension of the problem, which, especially for large finite-horizon problems, may lead to substantially speedier computation. Fourth, rather than solving the problem of singular transition and control-cost matrices by increasing the dimension of the state vector as is the common approach in recent work, our procedure reduces the dimensionality. Fifth, no eigenvalue computations or Jordan or Schur decompositions, with possible accompanying

numerical instabilities, are required for our approach.

Aside from procedural advantages, the results in the paper provide direct analytical insights. For instance, the fact in Theorem 2 that, when B_2 has full rank, the optimal feedback matrix can be obtained as a linear transformation of the dynamic core of the Riccati matrix may be valuable in theoretical applications. Some of the results developed in this paper appear to be quite novel from an applied mathematical perspective and may turn out to be of use in establishing further results and in other applications. For instance, the rank of the B_2 matrix may be helpful in dealing with identification problems in the Kalman context. Or, some of the results may be applicable to other issues in least squares-based time-series econometrics.

Appendix A: Proof of Theorem 1*

Consider a generalized inverse of B_2 given as: $B_2^I = S I_r^q Q$. It is easy to check that B_2^I is indeed a generalized inverse of B_2 since $B_2 B_2^I B_2 = Q^{-1} (I_r^q)^3 S^{-1} = B_2$ and since I_r^q is an idempotent matrix. Define $\Theta = (I_q - B_2^I B_2) S \begin{bmatrix} 0_{r, q-r} \\ I_{q-r} \end{bmatrix}$. Then by design $B_2 \Theta = 0_{q, q-r}$ and

straightforward multiplication shows that:

$$(A1) \quad \Theta = [I_q - S (I_r^q)^2 S^{-1}] S \begin{bmatrix} 0_{r, q-r} \\ I_{q-r} \end{bmatrix} = S (I_q - I_r^q) \begin{bmatrix} 0_{r, q-r} \\ I_{q-r} \end{bmatrix} = \begin{bmatrix} S_2 \\ S_4 \end{bmatrix}.$$

Since $B_2 \Theta = 0$ it follows from equation (11) updated one period that:

$$(A2) \quad (\Phi_t^{-1} - B_1) \Theta = (\Phi_t^{-1} - B_1) \begin{bmatrix} S_2 \\ S_4 \end{bmatrix} = 0_{q, q-r}, \quad t \leq T - 1.$$

Partition Φ_t^{-1} and B_1 according to S and extract the $r \times r$ upper left block of Φ_t^{-1} as $Z' \Phi_t^{-1} Z$.

Take S_4 as invertible⁴ and solve equations (A2) similarly to equation (6). This yields:

⁴ If S_4 is not invertible we can always rearrange the rows of S by pre-multiplying by some permutation matrix G such that $\hat{S} \equiv G S$ with \hat{S}_4 invertible. In this case transform all B_i to obtain $G' B_i G$. In order for equation (11) to continue to hold we must also transform Φ_t^{-1} to $G' \Phi_t^{-1} G$ and in order for the procedure to work we must also set $\hat{Q} \equiv Q G'$. Then we can proceed as above. We then obtain $G' (\Phi_t^{-1} - B_1) G$ in equation (A3) but can recover Φ_t^{-1} by inverting the known G matrix.

$$(A3) \quad \Phi_t^{-1} - B_1 = M^* (\Phi_t^{*-1} - B_{11}) M^{*'}, \quad M^* = \begin{bmatrix} I_r \\ -S_4^{-1'} S_2' \end{bmatrix}, \quad t \leq T-1,$$

where $\Phi_t^{*-1} \equiv Z' \Phi_t^{-1} Z$ represents the upper left block of Φ_t^{-1} and $B_{11} \equiv Z' B_1 Z$

represents the upper left block of B_1 . As Φ_t^{-1} and B_1 are positive definite, so are Φ_t^{*-1} and B_{11} .

To obtain the dynamics of $\Phi_t^{*-1} \equiv Z' \Phi_t^{-1} Z$, post-multiply equation (11) by Z and pre-multiply by Z' , and use equation (A3) on the right side to produce:

$$(A4) \quad \Phi_{t-1}^{*-1} - B_{11} = -Z' B_2' [M^* (\Phi_t^{*-1} - B_{11}) M^{*'} + (B_1 + B_3)]^{-1} B_2 Z, \quad t \leq T-1.$$

To manipulate equation (A4), consider that the term in brackets is positive definite (as it equals the sum of a positive definite matrix Φ_t^{-1} and a positive semi-definite matrix B_3). Further, $B_2 Z$

has full column rank r : $B_2 Z = Q^{-1} I_r^q S^{-1} Z$, $S^{-1} \equiv \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_3 & \Sigma_4 \end{bmatrix}$. Thus $I_r^q S^{-1} Z = \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix}$,

where $\text{rank} \left(\begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} \right) = \text{rank}(\Sigma_1)$, so $\text{rank}(B_2 Z) = \text{rank}(\Sigma_1)$. From footnote 4 we can

assume without loss of generality that S_4 has full rank $q - r$. We know that

$\Sigma_1 = (S_1 - S_2 S_4^{-1} S_3)^{-1}$ [Anderson and Moore (1990), p. 349] because S_4^{-1} exists and because

$|S_4| \cdot |S_1 - S_2 S_4^{-1} S_3| = |S| \neq 0$ so $|S_1 - S_2 S_4^{-1} S_3| \neq 0$ [Söderström (1994), p. 162]. Thus

Σ_1 has full rank r , so that $B_2 Z$ has full column rank. This fact, together with the positive definiteness of the term in brackets, establishes that the left-hand side of equation (A4) is negative definite and thus $\Phi_t^{*-1} - B_{11}$ is invertible. We can now use the inversion identity (see footnote

3) to rewrite the term in brackets, since the relevant inverses exist:

$$(A5) \quad \Phi_{t-1}^{*-1} = Z'[B_1 - B_2'(B_1 + B_3)^{-1}B_2]Z + Z'B_2'(B_1 + B_3)^{-1}M^* \cdot$$

$$[M^{*'}(B_1 + B_3)^{-1}M^* + (\Phi_t^{*-1} - B_{11})^{-1}]^{-1}M^{*'}(B_1 + B_3)^{-1}B_2Z, \quad t \leq T-1.$$

Again use the inversion identity to reformulate the second expression in brackets:

$$(A6) \quad \Phi_{t-1}^{*-1} = Z'[B_1 - B_2'(B_1 + B_3)^{-1}B_2]Z + Z'B_2'(B_1 + B_3)^{-1}M^* \left\{ [M^{*'}(B_1 + B_3)^{-1}M^*]^{-1} \right. \\ \left. - [M^{*'}(B_1 + B_3)^{-1}M^*]^{-1} \{ \Phi_t^{*-1} - Z'B_1Z + [M^{*'}(B_1 + B_3)^{-1}M^*]^{-1} \}^{-1} \right. \\ \left. [M^{*'}(B_1 + B_3)^{-1}M^*]^{-1} \right\} M^{*'}(B_1 + B_3)^{-1}B_2Z, \quad t \leq T-1.$$

Note that the inverse of the term in small braces exists by the invertibility of the second term in brackets in equation (A5), since, when $[X + Y^{-1}]^{-1}$ exists, so does $[Y + X^{-1}]^{-1}$ with

$$Y \equiv \Phi_t^{*-1} - B_{11} \text{ and } X \equiv M^{*'}(B_1 + B_3)^{-1}M^* \text{ both invertible.}$$

Equation (A6) directly yields equation (27) in Theorem 1*. The initial condition Φ_{T-1}^{*-1} is

obtained from using equation (11) in $\Phi_{T-1}^{*-1} \equiv Z'\Phi_{T-1}^{-1}Z$. ■

An example in which B_2^* is singular (after a single application of the reduction in Theorem

1*) is as follows:

$$K = I_4, \quad C = \begin{pmatrix} 0_{2,2} \\ I_2 \end{pmatrix}, \quad A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 2 \\ -0.5 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then $B_2 = A_1$ and so $\text{rank}(B_2) = 1$ and hence B_2^* is a scalar. Then

$$Q = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & -0.5 \\ 0 & -1 \end{pmatrix}, \quad \text{so } M^* = \begin{pmatrix} 1 \\ -0.5 \end{pmatrix}. \quad \text{Therefore } M^{*'}(B_1 + B_3)^{-1}B_2Z = 0$$

so that $B_2^* = 0$; hence $\text{rank } B_2^* = 0 < 1$.

Appendix B: Proof of Theorem 3*

First note that B^* exists since B_2^* is invertible after repeated application of Theorem 1*. Note

that since B_2^* is invertible we can apply to equation (27) the same transformations that were

applied to equation (11) in obtaining equation (25), and the result can be solved out to infer:

$$\Phi_{t-1}^* = [B_2^{*-1} + B_2^{*-1}B_3^*\Phi_t^*][B_1^*B_2^{*-1} + (B_1^*B_2^{*-1}B_3^* - B_2^{*'})\Phi_t^*]^{-1}, \quad t \leq T.$$

Then the proof proceeds as the proof of Theorem 3, with $N_T^* = \Gamma$ and $D_T^* = Z'M'K_T^{-1}MZ\Gamma$

for arbitrary invertible matrix Γ (as we know by applying $\Phi_t^{*-1} \equiv Z'\Phi_t^{-1}Z$ with

$\Phi_T^{-1} = M'K_T^{-1}M$) and with all Φ_t and B_i having asterisks attached. ■

References

- Amman, H. M., 1997, Numerical methods for linear-quadratic models, in H.M. Amman, D.A. Kendrick, and J. Rust, eds., *Handbook of Computational Economics* (North-Holland, Amsterdam).
- Amman, H. M. and H. Neudecker, 1997. Numerical solutions of the algebraic matrix Riccati equation, *Journal of Economic Dynamics and Control* 21, 363-369.
- Anderson, B. D. A. and J. B. Moore, 1979. *Optimal filtering* (Prentice Hall, Englewood Cliffs).
- Anderson, B. D. A. and J. B. Moore, 1990. *Optimal control: linear quadratic methods* (Prentice Hall, Englewood Cliffs).
- Anderson, E., L. P. Hansen, E. R. McGrattan and T. Sargent, 1997. Mechanics of forming and estimating dynamical linear economies, in H.M. Amman, D.A. Kendrick, and J. Rust, eds., *Handbook of Computational Economics* (North-Holland, Amsterdam).
- Anderson, G. and G. Moore, 1985. A linear algebraic procedure for solving linear perfect foresight models, *Economics Letters* 17, 247-252.
- Balvers, R. J. and T. F. Cosimano, 1994. Inflation variability and gradualist monetary policy, *Review of Economic Studies* 61, 721-738.
- Balvers, R. J. and Y. Wu, 2000. Momentum and mean reversion across national equity markets, Working paper, Rutgers University.
- Binder, M. and M. H. Pesaran, 2000. Solution of finite-horizon multivariate linear rational expectations models and sparse linear systems, *Journal of Economic Dynamics and Control* 24, 325-346.
- Blanchard, O. J. and Kahn, C. M., 1980. The solution of linear difference models under rational expectations, *Econometrica* 48, 1305-11.
- Chow, G. C., 1975. *Analysis and control of dynamic economic systems* (Wiley, New York).
- Claar, V. V., 2000. *The natural rate of unemployment, its estimation, and evaluation of its estimators*, Ph. D. dissertation, West Virginia University.
- Ehlgren, J., 1999. A nonrecursive solution method for the linear-quadratic optimal control problem with a singular transition matrix, *Computational Economics* 13, 17-23.

- Hannan, E. J. and M. Deistler, 1988. *The statistical theory of linear systems* (John Wiley and Sons, New York).
- Harvey, A. W., 1989. *Forecasting, structural time series models and the Kalman filter* (Cambridge University Press, New York).
- Kalman, R. E., 1960. A new approach to linear filtering and prediction problems, *Journal of Basic Engineering*, 35-45.
- King, R. G. and M. Watson, 1997. System reduction and solution algorithms for singular linear difference systems under rational expectations, Working paper, Princeton University.
- King, R. G. and M. Watson, 1998. The solution of singular linear-difference systems under rational expectations, *International Economic Review* 39, 1015-27.
- Klein, P., 2000. Using the generalized Schur form to solve a multivariate linear rational expectations model, *Journal of Economic Dynamics and Control* 24, 1405-24.
- Lancaster, P. and L. Rodman, 1995. *Algebraic Riccati Equations* (Clarendon Press, Oxford).
- Ljungqvist, L. and T. J. Sargent, 2000. *Recursive macroeconomic theory* (MIT Press, Cambridge).
- Martin, C. F., and G. Ammar, 1991. The geometry of the matrix Riccati equation and associated eigenvalue method, in Bittani, Laub, and Willems, eds., *The Riccati equation* (Springer Verlag, New York).
- Mitchell, D.W., 2000. An analytic Riccati solution for two-target discrete-time control, *Journal of Economic Dynamics and Control* 24, 615-622.
- Sims, C. S., 2000, Solving linear rational expectations models, Working paper, Yale University.
- Söderström, T., 1994. *Discrete-time stochastic systems: estimation and control* (Prentice-Hall, New York).
- Vaughan, D.R., 1970. A nonrecursive algebraic solution for the discrete Riccati equation, *IEEE Transactions on Automatic Control* 15, 597-599.

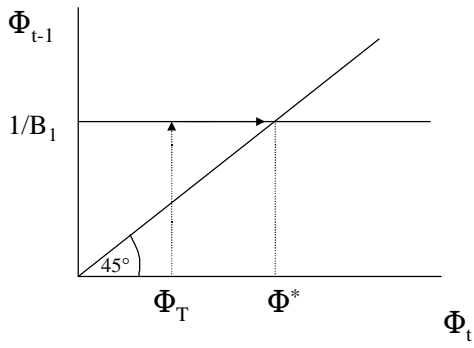


Fig. 1(a): $B_2 = 0$

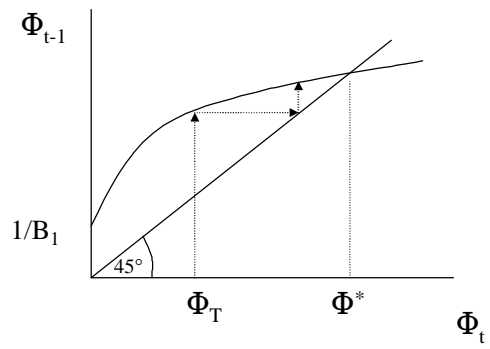


Fig. 1(b): $B_2 \neq 0, B_1 B_3 > B_2^2$

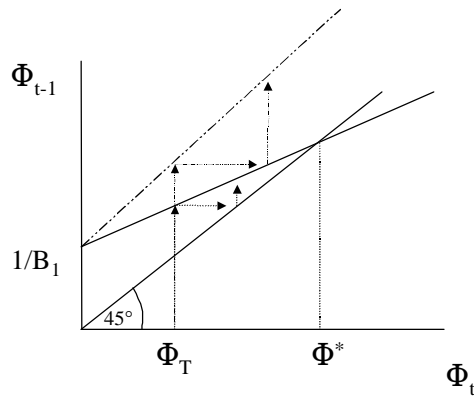


Fig. 1(c): $B_2 \neq 0, B_1 B_3 = B_2^2$