

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**A Reliable and Computationally Efficient Algorithm for
Imposing the Saddle Point Property in Dynamic Models**

Gary S. Anderson

2010-13

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

A Reliable and Computationally Efficient Algorithm for Imposing the Saddle Point Property in Dynamic Models

Gary S. Anderson

Abstract

This paper describes a set of algorithms for quickly and reliably solving linear rational expectations models. The utility, reliability and speed of these algorithms are a consequence of 1) the algorithm for computing the minimal dimension state space transition matrix for models with arbitrary numbers of lags or leads, 2) the availability of a simple modeling language for characterizing a linear model and 3) the use of the QR Decomposition and Arnoldi type eigenspace calculations. The paper also presents new formulae for computing and manipulating solutions for arbitrary exogenous processes.

1 Introduction and Summary

Economists at the Board have an operational need for tools that are useful for building, estimating and simulating moderate to large scale rational expectations models. This context dictates a need for careful attention to computational efficiency and numerical stability of the algorithms.

These algorithms have proved very durable and useful for staff at the central bank. Many economists at the Federal Reserve Board have used the algorithms in their daily work and their research.² With the exception of researchers at other European central banks(Zagaglia, 2002), few economists outside the US

¹ First, I would like to thank George Moore, my now deceased mentor, friend and coauthor of(Anderson & Moore, 1985). I also wish to thank Brian Madigan, Robert Tetlow, Andrew Levin, Jeff Fuhrer and Hoyt Bleakley for helpful comments. I am responsible for any remaining errors. The views expressed herein are mine and do not necessarily represent the views of the Board of Governors of the Federal Reserve System.

² See, for example,(Bomfim, 1996; Fuhrer & Moore, 1995a; Fuhrer & Moore, 1995b; Fuhrer & Moore, 1995; Fuhrer & Madigan, 1997; Fuhrer, 1997b; Fuhrer, 1997a;

central bank seem to know about the method. This paper attempts to make the method and approach more widely available by describing the underlying theory along with a number of improvements on the original algorithm.³

The most distinctive features of this approach are:

- its algorithm for computing the minimal dimension state space transition matrix
- its use of bi-orthogonality to characterize the asymptotic constraints that guarantee stability (See Section 3.1.2).
- It's reliance on QR Decomposition and the real Schur Decomposition for speed and accuracy.

This unique combination of features makes the algorithm especially effective for large models. See (Anderson, 2006) for a systematic comparison of this algorithm with the alternatives procedures.

The remainder of the paper is organized as follows. Section 2 states the saddle point problem. Section 3 describes the algorithms for solving the homogeneous and inhomogeneous versions of the problem and describes several implementations. Section 4 shows how to compute matrices often found useful for manipulating rational expectations model solutions: the observable structure(Section 4.1) and stochastic transition matrices(Section 4.2). The Appendices contain proofs for the linear algebra underlying the algorithm and the solution of a simple example model.

2 Saddle Point Problem Statement

Consider linear models of the form:

$$\sum_{i=-\tau}^{\theta} H_i x_{t+i} = \Psi z_t, \quad t = 0, \dots, \infty \quad (1)$$

Fuhrer, 1996; Fuhrer *et al.*, 1995; Fuhrer & Hooker, 1993; Fuhrer, 1994; Fuhrer, 1997c; Orphanides *et al.*, 1997; Levin *et al.*, 1998; Orphanides, 1998; Orphanides & Wieland, 1998; Edge *et al.*, 2003; Orphanides & Williams, 2002).

³ At the Board, economists commonly refer to this family of algorithms as the AIM algorithm. A metaphor relating our approach to the “shooting method” inspired the name.

with initial conditions, if any, given by constraints of the form

$$x_i = x_i^{data}, i = -\tau, \dots, -1 \quad (2)$$

where both τ and θ are non-negative, and x_t is an L dimensional vector of endogenous variables with

$$\lim_{t \rightarrow \infty} \|x_t\| < \infty \quad (3)$$

and z_t is a k dimensional vector of exogenous variables.

Section 3 describes computationally efficient algorithms for determining the existence and uniqueness of solutions to this problem.

3 The Algorithms

The uniqueness of solutions to system 1 requires that any transition matrix characterizing the dynamics of the linear system have an appropriate number of explosive and stable eigenvalues (Blanchard & Kahn, 1980), and that a certain set of asymptotic linear constraints are linearly independent of explicit and certain other auxiliary initial conditions (Anderson & Moore, 1985).

The solution methodology entails

- (1) Manipulating the left hand side of equation 1 to obtain a state space transition matrix, A , along with a set of auxiliary initial conditions, Z for the homogeneous solution.

$$Z \begin{bmatrix} x_{-\tau} \\ \vdots \\ x_\theta \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} x_{-\tau+1} \\ \vdots \\ x_\theta \end{bmatrix} = A \begin{bmatrix} x_{-\tau} \\ \vdots \\ x_{\theta-1} \end{bmatrix} \quad (4)$$

See Section 3.1.1.

- (2) Computing the eigenvalues and vectors spanning the left invariant space associated with large eigenvalues. See Section 3.1.2.

$$VA = \mathcal{M}V \quad (5)$$

with the eigenvalues of \mathcal{M} all greater than one in absolute value.

- (3) Assembling asymptotic constraints, Q , (See Section 3.1.2.) by combining the:
 - (a) auxiliary initial conditions identified in the computation of the transition matrix and

(b) the invariant space vectors

$$Q = \begin{bmatrix} Z \\ V \end{bmatrix} \quad (6)$$

- (4) Investigating the rank of the linear space spanned by these asymptotic constraints and, when a unique solution exists,
- (a) Computing the auto-regressive representation, B . See Section 3.1.3.
 - (b) Computing matrices, ϕ, F, ϑ for characterizing the impact of the inhomogeneous right hand side term. See Section 3.2.1.

Figure 1 presents a flowchart of the algorithm.

3.1 Homogeneous Solution

Suppose, for now, that $\Psi = 0$:⁴

$$\sum_{i=-\tau}^{\theta} H_i x_{t+i} = 0, t \geq 0 \quad (8)$$

$$\lim_{t \rightarrow \infty} \|x_t\| < \infty \quad (9)$$

The homogeneous specification 8 is not restrictive. Since the procedure can handle inhomogeneous versions of equation 1 by recasting the problem in terms of deviations from a steady state value. However, the next section provides a more intuitive, flexible and computationally efficient alternative for computing inhomogeneous solutions.⁵

⁴ Note that there is no unique steady state requirement. Steady state solutions, x^* satisfying

$$\sum_{i=-\tau}^{\theta} (H_i) x^* = 0 \quad (7)$$

lie in a linear subspace of \mathcal{R}^L . We will develop conditions that guarantee solutions that evolve from a given set of initial conditions to a single point in this subspace. As a result, one can apply these routines to models with unit roots, seasonal factors, cointegrating vectors and error correction terms.

⁵ The original algorithmic description and software implementation of these algorithms developed homogeneous solutions. Researchers obtained solutions for models with inhomogeneous systems by adding an equation of the form $\mathbf{i}_t = \mathbf{i}_{t-1}$ with initial condition $\mathbf{i}_{t-1} = 1$ to the system.

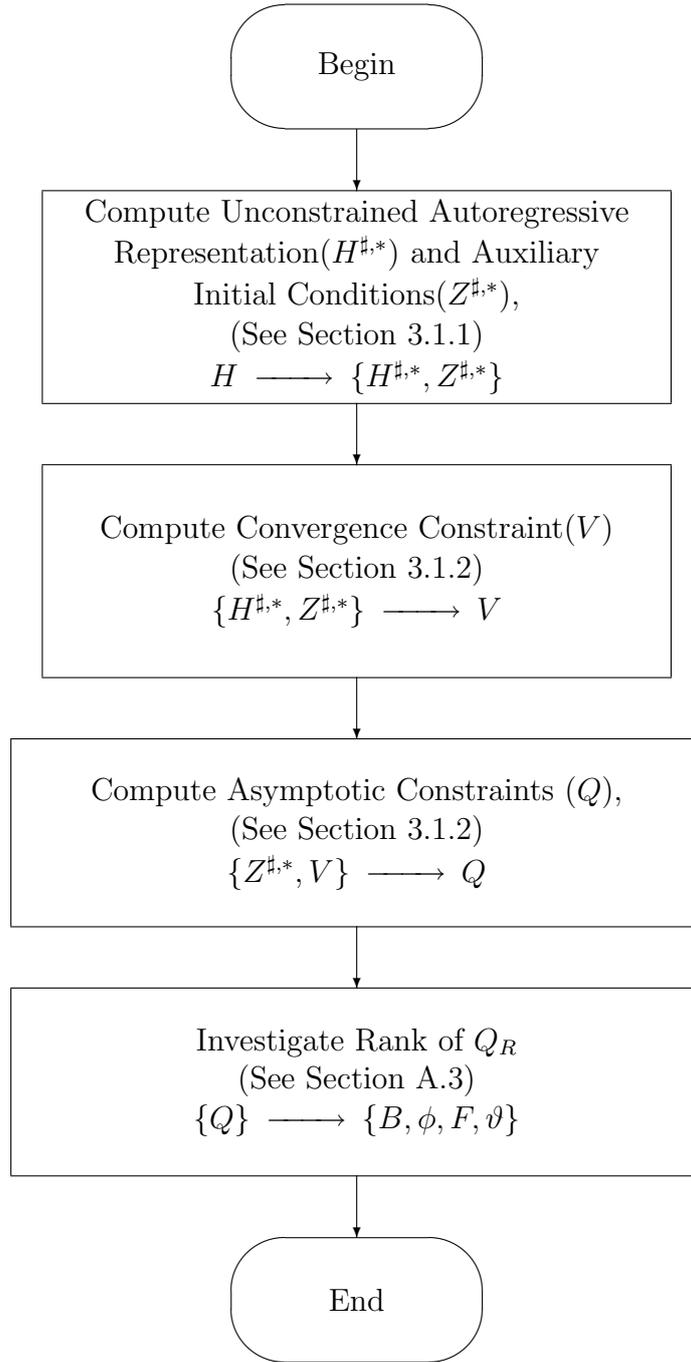


Fig. 1. Algorithm Components

3.1.1 State Space Transition Matrix and Auxiliary Initial Conditions: A, Z

This section describes how to determine a first-order state space representation of the equation system 8. The method is an extension of the shuffling algorithm developed in (Luenberger, 1978; Luenberger, 1977). If H_θ is non-singular, we

can immediately obtain $x_{t+\theta}$ in terms of $x_{t-\tau} \dots x_{t+\theta-1}$

$$-H_\theta^{-1} \begin{bmatrix} H_{-\tau} & \dots & H_{\theta-1} \end{bmatrix} \quad (10)$$

However, the natural specification of many economic models has singular H_θ .

This, first, algorithm applies full rank linear transformations to equations from the original linear system in order to express $x_{t+\theta}$ in terms of $x_{t-\tau} \dots x_{t+\theta-1}$. It produces an unconstrained, typically explosive, autoregressive representation for the evolution of the components of the state space vectors and a set of vectors that provide important auxiliary initial conditions.

Section A.1 presents a proof that repeating this process of annihilating and regrouping rows ultimately produces an $H^{\#,k} = H^{\#,*}$ with $H_\theta^{\#,*}$ non-singular. The proof identifies a benign rank condition that guarantees that the algorithm will successfully compute the unconstrained autoregression and the auxiliary initial conditions.

Figures 2-4 provide a graphical characterization of the linear algebraic transformations characterizing the algorithm. Figure 2 presents a graphical characterization of the relevant set of linear constraints for $t = 0 \dots (\tau + \theta)$. The figure represents the regions where the coefficients are potentially non-zero as shaded gray. If $H_\theta^{\#,0}$ is singular, one can find a linear combination of its rows which preserves the rank of $H_\theta^{\#,0}$ but which annihilates one or more of its rows.

Consequently, one can pre-multiply the matrices presented in Figure 2

by a unitary matrix to get the distribution of zeros displayed in Figure 3. Since the matrices repeat over time, one need only investigate the rank of the square matrix $H_\theta^{\#,0}$. $U_1 H_\theta^{\#,0} =$

Fig. 2. Initial Tableau

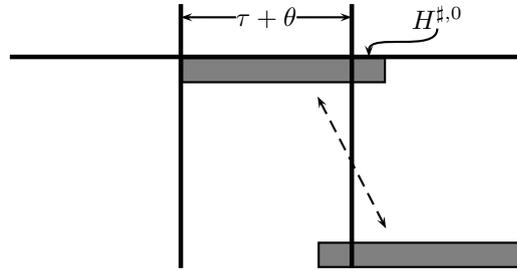


Fig. 3. Forward Row Annihilation

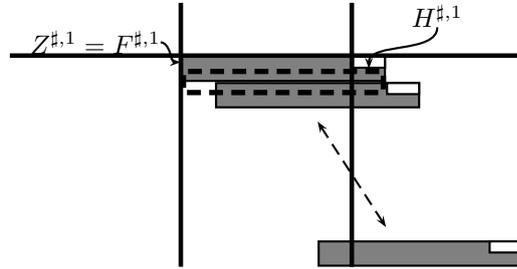
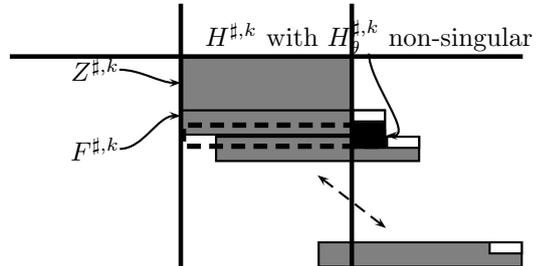


Fig. 4. Full Rank Leading Block



$\left[F^{\sharp,1} U_1 H_\theta^{\sharp,0} \right]$. With some of the rows of $U_1 H_\theta^{\sharp,0}$ all zeros. One can regroup the rows in the new tableau to get $H^{\sharp,1}$. By construction, the rank of $H_\theta^{\sharp,1}$ can be no less than the rank of $H_\theta^{\sharp,0}$.

Note that changing the order of the equations in $H_\theta^{\sharp,0}$ will not affect the rank of $H_\theta^{\sharp,0}$ or the space spanned by the nonzero rows that will enter $H_\theta^{\sharp,1}$. Since $H_\theta^{\sharp,0}$ is not full rank, a subset of the rows of $H_\theta^{\sharp,1}$ will span the same space as $H_\theta^{\sharp,0}$. The other rows in $H_\theta^{\sharp,1}$ will come from linear combinations of the original system of equations in effect “shifted forward” in time.

One can think of the “Full Rank Leading Block” matrix as the result of pre-multiplications of the “Initial Tableau” by a sequence of unitary matrices. Implementations of the algorithm can take advantage of the fact that the rows of the matrices repeat. The regrouping can be done by “shifting equations forward” in time in an $L \times L(\tau + 1 + \theta)$ version of the tableau.

Section A.1 presents a proof that repeating this process of annihilating and regrouping rows ultimately produces an $H^{\sharp,k} = H^{\sharp,*}$ with $H_\theta^{\sharp,*}$ non-singular. Algorithm 1 presents pseudo code for an algorithm for computing the components of the state space transition matrix and the auxiliary initial conditions.

Algorithm 1

```

1  Given  $H$ ,
2  compute the unconstrained autoregression.
3  funct  $\mathcal{F}_1(H) \equiv$ 
4     $k := 0$ 
5     $\mathcal{Z}^0 := \emptyset$ 
6     $\mathcal{H}^0 := H$ 
7     $\Gamma = \emptyset$ 
8    while  $\mathcal{H}_\theta^k$  is singular  $\cap \text{rows}(\mathcal{Z}^k) < L(\tau + \theta)$ 
9      do
10     Determine a Non-singular matrix that annihilates  $L - r(\mathcal{H}_\theta^k)$  Rows of  $\mathcal{H}_\theta^k$ 
11      $U^k = \begin{bmatrix} U_Z^k \\ U_N^k \end{bmatrix} := \text{rowAnnihilator}(\mathcal{H}_\theta^k)$ 
12      $\mathcal{H}^{k+1} := \begin{bmatrix} 0 & U_Z^k \mathcal{H}_\tau^k & \dots & U_Z^k \mathcal{H}_{\theta-1}^k \\ U_N^k \mathcal{H}_\tau^k & \dots & U_N^k \mathcal{H}_\theta^k \end{bmatrix}$ 
13      $\mathcal{Z}^{k+1} := \begin{bmatrix} \mathcal{Z}^k \\ U_Z^k \mathcal{H}_\tau^k \dots U_Z^k \mathcal{H}_{\theta-1}^k \end{bmatrix}$ 
14      $k := k + 1$ 
15   od

```

```

16   $\Gamma = -H_\theta^{-1} \begin{bmatrix} H_{-\tau} & \dots & H_{\theta-1} \end{bmatrix}$ 
17   $A = \begin{bmatrix} 0 & I \\ \Gamma & \end{bmatrix}$ 
18  return  $\{ \begin{bmatrix} \mathcal{H}_{-\tau}^k & \dots & \mathcal{H}_\theta^k \end{bmatrix}, A, \mathcal{Z}^k \}$ 
19  .

```

The algorithm terminates with:

$$\begin{bmatrix} H_{-\tau}^{\sharp*} & \dots & H_\theta^{\sharp*} \end{bmatrix} \begin{bmatrix} x_{-\tau} \\ \vdots \\ x_\theta \end{bmatrix} = 0 \quad (11)$$

with $H_\theta^{\sharp*}$ non singular. Let

$$\Gamma^\sharp = -(H_\theta^{\sharp*})^{-1} \begin{bmatrix} H_{-\tau}^{\sharp*} & \dots & H_{\theta-1}^{\sharp*} \end{bmatrix} \quad (12)$$

Then

$$x_\theta = \Gamma^\sharp \begin{bmatrix} x_{-\tau} \\ \vdots \\ x_{\theta-1} \end{bmatrix} \quad (13)$$

This unconstrained auto-regression in x_t provides exactly what one needs to construct the state space transition matrix.

$$A^\sharp = \begin{bmatrix} 0 & I \\ \Gamma^\sharp & \end{bmatrix} \quad (14)$$

so that

$$\begin{bmatrix} x_{-\tau+1} \\ \vdots \\ x_\theta \end{bmatrix} = A^\sharp \begin{bmatrix} x_{-\tau} \\ \vdots \\ x_{\theta-1} \end{bmatrix} \quad (15)$$

3.1.2 Asymptotic Linear Constraint Matrix: Q

In order to compute solutions to equation 8 that converge, one must rule out explosive trajectories. Blanchard and Kahn (Blanchard & Kahn, 1980) used eigenvalue and eigenvector calculations to characterize the space in which the solutions must lie. In contrast, our approach uses an orthogonality constraint to characterize regions which the solutions must avoid.

Each left eigenvector associated with a given eigenvalue is orthogonal to each right eigenvector associated with roots associated with different eigenvalues. Since vectors in the left invariant space associated with roots outside the unit circle are orthogonal to right eigenvectors associated with roots inside the unit circle, a given state vector that is part of a convergent trajectory must be orthogonal to each of these left invariant space vectors. See theorem 4 on page 20. Thus, the algorithm can exploit bi-orthogonality and a less burdensome computation of vectors spanning the left invariant space in order to rule out explosive trajectories.

If the vectors in V span the invariant space associated with explosive roots, trajectories satisfying equation 8 are non-explosive if and only if

$$VA = \mathcal{M}V \tag{16}$$

with the eigenvalues of \mathcal{M} similar to the Jordan blocks of A associated with all eigenvalues greater than one in absolute value.

$$V \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t+\theta-1} \end{bmatrix} = 0 \tag{17}$$

for some t .⁶

Combining V and $Z^\#$ completely characterizes the space of stable solutions satisfying the linear system 8.

$$Q = \begin{bmatrix} Z^\# \\ V \end{bmatrix} \tag{18}$$

⁶ If A has roots with magnitude 1 then trajectories can converge to either a limit cycle or a non-zero fixed point. Otherwise, non-explosive trajectories will converge to the origin.

The first set of equations come from the equations in equation system 8 which do not appear in the transformed system of Equation 11 but must nevertheless be satisfied. The second set of equations come from the constraint that the solutions are non-explosive. Algorithm 2 provides pseudo code for computing Q .

Algorithm 2

```

1 Given  $A, Z^{\#, *}$ ,
2 func  $\mathcal{F}_2(A, Z^{\#, *})$ 
3   Compute  $V$ , the vectors spanning the left
4   invariant space of  $A$  associated with eigenvalues
5   greater than one in magnitude
6    $Q := \begin{bmatrix} Z^{\#, *} \\ V \end{bmatrix}$ 
7   return  $Q$ 
8 .

```

3.1.3 Convergent Autoregressive Representation: B

The first two algorithms together produce a matrix Q characterizing constraints guaranteeing that trajectories are not explosive. See theorem 5 and corollary 2 for a proof. (Hallet & McAdam, 1999) describes how to use the matrix Q from Section 3.1.2 to impose saddle point stability in non linear perfect foresight models. However, for linear models with unique saddle point solutions it is often useful to employ an autoregressive representation of the solution. Theorem 6 in Section A.3 provides a fully general characterization of the existence and uniqueness of a saddle point solution.

A summary for typical applications of the algorithm follows. Partition $Q = \begin{bmatrix} Q_L & Q_R \end{bmatrix}$ where Q_L has $L\tau$ columns. When $\eta = L\theta$, Q_R is square. If Q_R is

non-singular, the system has a unique solution⁷

$$\begin{bmatrix} B \\ B \\ 2 \\ \vdots \\ B \\ \theta \end{bmatrix} = Q_R^{-1}Q_L \quad \text{and solutions are of the form } x_t = B \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix}, x_{t+k} = \frac{B}{k} \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix} \quad (19)$$

Algorithm 3 provides pseudo code for computing B.

Algorithm 3

```

1 Given Q,
2 funct  $\mathcal{F}_3(Q)$ 
3   cnt = noRows(Q)
4   return  $\begin{cases} \{Q, \infty\} & \text{cnt} < L\theta \\ \{Q, 0\} & \text{cnt} > L\theta \\ \{Q, \infty\} & (Q_R \text{ singular}) \\ \{B = -Q_R^{-1}Q_L, 1\} & \text{otherwise} \end{cases}$ 
5 .

```

3.2 *Inhomogeneous Solution*

Now, suppose

$$\sum_{i=-\tau}^{\theta} H_i x_{t+i} = \Psi z_t, t \geq 0 \quad (20)$$

$$\lim_{t \rightarrow \infty} \|x_t\| < \infty \quad (21)$$

3.2.1 *Inhomogeneous Factor Matrices: ϕ, F*

Theorem 1 Given structural model matrices, $H_i, i = -\tau, \dots, \theta$ and Ψ , convergent autoregression matrix B there exist *inhomogeneous*

⁷ If Q_R is singular, the system has an infinity of solutions. When $\eta < L\theta$, The system has an infinity of solutions.

When Q has more than $L\theta$ rows, The system has a unique nontrivial solution only for specific values of x_{data} .

3.2.2 Exogenous VAR Impact Matrix, ϑ

Modelers can augment the homogeneous linear perfect foresight solutions with particular solutions characterizing the impact of exogenous vector autoregressive variables.

Theorem 2 When

$$z_{t+1} = \Upsilon z_t \quad (25)$$

one can show that

$$x_t = \begin{bmatrix} B_L & B_R \end{bmatrix} \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix} + \vartheta z_t \quad (26)$$

where

$$vec(\vartheta) = \begin{bmatrix} 0 & \dots & 0 & I \end{bmatrix} (I - \Upsilon^T \otimes F)^{-1} vec \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \phi\Psi \end{bmatrix} \quad (27)$$

See Section A.5 for derivations and formulae.

3.3 Implementations

This set of algorithms has been implemented in a wide variety of languages. Three implementations, a Matlab, a “C”, and a Mathematica implementation, are available from the author. Each implementation avoids using the large tableau of Figure 2. They each shift elements in the rows of a single copy of the matrix H . Each implementation eliminates inessential lags from the autoregressive representation before constructing the state space transition matrix for invariant space calculations.

The most widely used version is written in MATLAB. The MATLAB version has a convenient modeling language front end for specifying the model equations and generating the H_i matrices.

The “C” version, designed especially for solving large scale models is by far the fastest implementation and most frugal with memory. It uses sparse lin-

ear algebra routines from SPARSKIT(Saad, 1994) and HARWELL(Numerical Analysis Group, 1995) to economize on memory. It avoids costly eigenvector calculations by computing vectors spanning the left invariant space using ARPACK(Lehoucq *et al.*, 1996).

For small models, one can employ a symbolic algebra version of the algorithms written in Mathematica. On present day computers this, code can easily construct symbolic state space transition matrices and compute symbolic expressions for eigenvalues for models with 5 to 10 equations. The code can often obtain symbolic expressions for the invariant space vectors when the transition matrix is of dimension 10 or less.

4 Other Useful Rational Expectations Solution Calculations

Economists use linear rational expectations models in a wide array of application. The following sections describe calculations which are useful for optimal policy design, model simulation and estimation exercises.

4.1 Observable Structure: S

To compute the error term for estimation of the coefficients of these models, one must commit to a particular information set. Two typical choices are t and $t-1$ period expectations.

Given structural model matrices, $H_i, i = -\tau, \dots, \theta$ and convergent autoregression matrices $B_i, i = -\tau, -1$ there exists an observable structure matrix, S

$$\epsilon_t = S \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_t \end{bmatrix} \quad (28)$$

See Section A.6 for a derivation and formula for S . Algorithm 5 provides pseudo code for computing S for a given lag, k^* in the availability of information.

Algorithm 5


```

3    $\mathcal{A} = \begin{bmatrix} 0 & & & I \\ \vdots & & \ddots & \\ 0 & & & I \\ S_0^{-1}S_{t-\tau-\max(k^*-1,0)+1} & \dots & \dots & S_0^{-1}S_{-1} \end{bmatrix}$ 
4    $\mathcal{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_0^{-1} \end{bmatrix}$ 
5   return( $\mathcal{A}, \mathcal{B}$ )
6   .

```

5 Conclusions

This paper describes a set of algorithms that have proved very durable and useful for staff at the central bank. The most distinctive features of this approach are:

- its algorithm for computing the minimal dimension state space transition matrix
- its use of bi-orthogonality to characterize the asymptotic constraints that guarantee stability.
- It's reliance on QR Decomposition and the real Schur Decomposition for speed and accuracy.

The unique combination of features makes the algorithm more efficient than all the alternatives—especially for large models. Staff at the Federal Reserve have developed a large scale linear rational expectations model consisting of 421 equations with one lead and one lag. This model provides an extreme example of the speed advantage of the Anderson-Moore Algorithm(AMA). On an Intel(R) Xeon 2.80GHz CPU running Linux the MATLAB version of AMA computes the rational expectations solution in 21 seconds while the the MATLAB version of gensys, a popular alternative procedure, requires 16,953 seconds. See (Anderson, 2006) for a systematic comparison of this algorithm with the alternative procedures. The code is available for download at

<http://www.federalreserve.gov/Pubs/oss/oss4/aimindex.html>

References

- Anderson, Gary. 2006. *Solving Linear Rational Expectations Models: A Horse Race*. Tech. rept. Federal Reserve System, Finance and Economics Discussion Series.
- Anderson, Gary, & Moore, George. 1985. A Linear Algebraic Procedure For Solving Linear Perfect Foresight Models. *Economics Letters*, **17**(3).
- Blanchard, Olivier Jean, & Kahn, Charles M. 1980. The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, **48**(5), 1305–11.
- Bomfim, Antulio. 1996. *Forecasting the Forecasts of Others: Expectational Heterogeneity and Aggregate Dynamics*. Tech. rept. 1996-41. Federal Reserve Board, Finance and Economics Discussion Series.
- Edge, Rochelle M., Laubach, Thomas, & Williams, John C. 2003. *The responses of wages and prices to technology shocks*. Tech. rept. 2003-21. Federal Reserve Bank of San Francisco, Working Papers in Applied Economic Theory.
- Fuhrer, Jeffrey C. 1994. Optimal Monetary Policy and the Sacrifice Ratio. *Pages 43–69 of: Fuhrer, Jeffrey C. (ed), 'Goals, Guidelines, and Constraints Facing Monetary Policymakers'*. Federal Reserve Bank of Boston Conference Series No. 38.
- Fuhrer, Jeffrey C. 1996. Monetary Policy Shifts and Long-term Interest Rates. *Quarterly Journal of Economics*, **111**(November), 1183–1209.
- Fuhrer, Jeffrey C. 1997a. Inflation/Output Variance Trade-offs and Optimal Monetary Policy. *Journal of Money Credit and Banking*, **29**, No. 2(May), 214–234.
- Fuhrer, Jeffrey C. 1997b. The (Un)Importance of Forward-Looking Behavior in Price Specifications. *Journal of Money Credit and Banking*, **29**, No. 3(August), 338–350.
- Fuhrer, Jeffrey C. 1997c. Towards a Compact, Empirically-Verified Rational Expectations Model for Monetary Policy Analysis. *Carnegie-Rochester Conference Series on Public Policy*, forthcoming.
- Fuhrer, Jeffrey C., & Hooker, Mark W. 1993. Learning About Monetary Regime Shifts in an Overlapping Wage Contract Economy. *Journal of Economic Dynamics and Control*, **17**, 531–553.
- Fuhrer, Jeffrey C., & Madigan, Brian. 1997. Monetary Policy When Interest Rates are Bounded at Zero. *Review of Economics and Statistics*, **79**, No. 4(November), 573–585.
- Fuhrer, Jeffrey C., & Moore, George R. 1995. Forward-Looking Behavior and the Stability of a Conventional Monetary Policy Rule. *Journal of Money Credit and Banking*, **27**, No. 4(November), 1060–70.
- Fuhrer, Jeffrey C., & Moore, George R. 1995a. Inflation Persistence. *Quarterly Journal of Economics*, **110**(February), 127–159.
- Fuhrer, Jeffrey C., & Moore, George R. 1995b. Monetary Policy Trade-Offs

- and the Correlation Between Nominal Interest Rates and Real Output. *American Economic Review*, **85**(March), 219–239.
- Fuhrer, Jeffrey C., Moore, George, & Schuh, Scott. 1995. A Comparison of GMM and Maximum-Likelihood Estimates of Quadratic Cost Inventory Models. *Journal of Monetary Economics*, **35**(February), 115–157.
- Hallet, Andrew Hughes, & McAdam, Peter (eds). 1999. *Analyses in Macroeconomic Modelling*. Kluwer Academic Publishers. Chap. Accelerating Non Linear Perfect Foresight Solution by Exploiting the Steady State Linearization.
- Lehoucq, R.B., Sorensen, D.C., & Yang, C. 1996. *ARPACK USERS GUIDE: Solution of Large Scale Eigenvalue Problems by Implicitly Restarted Arnoldi Methods*. Tech. rept. Department of Computational and Applied Mathematics Rice University.
- Levin, Andrew, Williams, John, & Wieland, Volker. 1998 (January). *Are Simple Rules Robust under Model Uncertainty?* Seminar Paper.
- Luenberger, David G. 1977. Dynamic Equations in Descriptor Form. *IEEE Transactions on Automatic Control*, 312–321.
- Luenberger, David G. 1978. Time-Invariant Descriptor Systems. *Automatica*, **14**, 473–480.
- Noble, Ben. 1969. *Applied Linear Algebra*. Prentice-Hall, Inc.
- Numerical Analysis Group. 1995. *The Harwell Sparse Matrix Library Manual*. Tech. rept. The Numerical Analysis Group.
- Orphanides, Athanasios. 1998. *Monetary Policy Evaluation with Noisy Information*. Tech. rept. 98-50. Finance and Economics Discussion Series.
- Orphanides, Athanasios, & Wieland, Volker. 1998 (January). *Price Stability and Monetary Policy Effectiveness when Nominal Interest Rates are Bounded at Zero*. Working Paper.
- Orphanides, Athanasios, & Williams, John C. 2002. Robust Monetary Policy Rules with Unknown Natural Rates. *Brookings Papers on Economic Activity*.
- Orphanides, Athanasios, Small, David, Wilcox, David, & Wieland, Volker. 1997 (May). *A Quantitative Exploration of the Opportunistic Approach to Disinflation*. Tech. rept. 97-36. Federal Reserve Board, Finance and Economics Discussion Series.
- Saad, Youcef. 1994. *SPARSKIT: A Basic Tool Kit for Sparse Matrix Computations*. Tech. rept. Computer Science Department, University of Minnesota.
- Taylor, John. 1979. Staggered Wage Setting in a Macro Model. *The American Economic Review*, **69**(2), 108–113.
- Zagaglia, Paolo. 2002. *Matlab Implementation of the AIM Algorithm: A Beginner's Guide*. Tech. rept. 169. Universita' Politecnica delle Marche (I), Dipartimento di Economia.

A Proofs

A.1 Unconstrained Autoregression

Theorem 3 Let

$$\mathcal{H} = \left[\begin{array}{cccc} H_{-\tau} & \dots & & H_{\theta} \\ & H_{-\tau} & \dots & H_{\theta} \\ & & \ddots & \\ & & & H_{-\tau} & \dots & H_{\theta} \end{array} \right] \left. \vphantom{\begin{array}{cccc} H_{-\tau} & \dots & & H_{\theta} \\ & H_{-\tau} & \dots & H_{\theta} \\ & & \ddots & \\ & & & H_{-\tau} & \dots & H_{\theta} \end{array}} \right\}^{\tau+\theta+1} \quad (\text{A.1})$$

There are two cases:

- When \mathcal{H} is full rank the algorithm terminates with $Z^{\#*}$ (Z^{b*}) and non-singular $H_{\theta}^{\#*}$ (H_{τ}^{b*})
- When \mathcal{H} is not full rank the algorithm terminates when some row of $[\mathcal{H}_{-\tau}^k \dots \mathcal{H}_{\theta}^k]$ is zero.

Proof Consider the case when \mathcal{H} is full rank. Each step of the algorithm applies a rank preserving pre-multiplication by a non singular matrix. Each step of the algorithm where $H_{\theta}^{\#,k}$ is singular, increases the row rank of $Z^{\#,k}$ by at least one. At each step $Z^{\#,k}$ are full rank. The rank of $Z^{\#,k}$ cannot exceed $L(\tau + \theta)$. \blacksquare

The following corollary indicates that models with unique steady-states always terminate with non singular $\mathcal{H}_{\theta}^{\#,*}$.

Corollary 1 If $(\sum_{i=-\tau}^{\theta} H_i)$ is non singular then

- (1) \mathcal{H} is full rank.
- (2) The origin is the unique steady state of equation 1.
- (3) there exists a sequence of elementary row operations that transforms \mathcal{H} into \mathcal{H}^*

Proof Suppose \mathcal{H} is not full rank. Then there is a non zero vector V such that $V\mathcal{H} = 0$. Consequently,

$$\left[V_{-\tau} \dots V_{\theta} \right] \mathcal{H} \begin{bmatrix} I \dots I \\ \vdots \dots \vdots \\ I \dots I \end{bmatrix} = 0 \quad (\text{A.2})$$

and

$$V_i \left(\sum_{j=-\tau}^{\theta} H_j \right) = 0 \forall i \quad (\text{A.3})$$

So that $(\sum_{i=-\tau}^{\theta} H_i)$ must be singular. ■

A.2 Asymptotic Constraints

Theorem 4 Consider a left invariant space and a right invariant space with no eigenvalues in common. Suppose V_1 spans the left invariant space and W_2 spans the right invariant space.

$$V_1 A = T_1 V_1 \quad (\text{A.4})$$

$$A W_2 = W_2 T_2 \quad (\text{A.5})$$

With eigenvalues of $T_1 \neq T_2$. Then $V_1 W_2 = 0$

Proof A right eigenvector x_i and a left-eigenvector y_j corresponding to distinct eigenvalues λ_i and λ_j are orthogonal. (Noble, 1969) Finite dimensional matrices have finite dimensional Jordan blocks. Raising a given matrix to a power produces a matrix with smaller Jordan blocks. Raising the matrix to a high enough power ultimately eliminates all nontrivial Jordan Blocks. Consequently, the left invariant space vectors are linear combination of the left eigenvectors and the right invariant space vectors are a linear combination of the right eigenvectors of the transition matrix raised to some finite power.

$$V_1 = \beta_1 \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix} \quad (\text{A.6})$$

$$W_2 = \begin{bmatrix} x_1 & \dots & x_K \end{bmatrix} \alpha_2 \quad (\text{A.7})$$

$$V_1 W_2 = \beta_1 \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix} \begin{bmatrix} x_1 & \dots & x_K \end{bmatrix} \alpha_2 = 0 \quad (\text{A.8})$$

■

Theorem 5 Let $\{x_t^{conv}\}$, $t = -\tau, \dots, \infty$ be a non explosive solution satisfying equation 1. Let A be the state space transition matrix for

equation 1 and V be a set of invariant space vectors spanning the invariant space associated with roots of A of magnitude bigger than 1. Then for $t = 0, \dots, \infty$

$$V \begin{bmatrix} x_{t-\tau}^{conv} \\ \vdots \\ x_{t+\theta-1}^{conv} \end{bmatrix} = 0 \quad (\text{A.9})$$

Proof Using W , the left generalized eigenvectors of A , one can employ the Jordan Canonical Form of A to write

$$W^H A = J W^H \quad (\text{A.10})$$

so that

$$A^t = (W^H)^{-1} J^t W^H \quad (\text{A.11})$$

$$y_t = A^t y_0 \quad (\text{A.12})$$

$$W^H y_t = J^t W^H y_0 \quad (\text{A.13})$$

$$\lim_{t \rightarrow \infty} y_t = 0 \Rightarrow \lim_{t \rightarrow \infty} W^H y_t = 0 \quad (\text{A.14})$$

Consequently,

$$W_i^H y_0 = 0 \forall i \text{ such that } |J_{ii}| > 1. \quad (\text{A.15})$$

so that

$$V y_0 = \alpha W^H y_0 = 0 \quad (\text{A.16})$$

■

Corollary 2 Let $\{x_t\}$, $t = -\tau, \dots, \infty$ be a solution satisfying equation 8. If A has no roots with magnitude 1 then the path converges to the origin if and only if

$$V \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t+\theta-1} \end{bmatrix} = 0 \quad (\text{A.17})$$

for some t .

Proof

$$W_i^H y_0 = 0 \forall i \text{ such that } |J_{ii}| > 1. \quad (\text{A.18})$$

means y_t

$$J_{ii} \neq 1. \quad (\text{A.19})$$

$$y_t = A^t y_0 \quad (\text{A.20})$$

■

A.3 Existence and Uniqueness

Theorem 6 Identify Q_L, Q_R from

$$Q = \begin{bmatrix} Z^\# \\ V \end{bmatrix} = \begin{bmatrix} Q_L & Q_R \end{bmatrix} \quad (\text{A.21})$$

with Q_R an $(\eta \times L\theta)$ matrix, Q_L an $(\eta \times L\tau)$ matrix, where η represent the number of rows in the matrix Q .

The existence of convergent solutions depends on the magnitudes of the ranks of the augmented matrix

$$r_1 = \text{rank} \left(\begin{bmatrix} Q_R & -Q_L x_{data} \end{bmatrix} \right) \quad (\text{A.22})$$

and

$$r_2 = \text{rank}(Q_R). \quad (\text{A.23})$$

By construction, $r_1 \geq r_2$ and $r_2 \leq L\theta$. There are three cases.

- (1) If $r_1 > r_2$ there is no nontrivial convergent solution
- (2) If $r_1 = r_2 = L\theta$ there is a unique convergent solution
- (3) If $r_1 = r_2 < L\theta$ the system has an infinity of convergent solutions

Corollary 3 When $\eta = L\theta$, Q_R is square. If Q_R is non-singular, the system has a unique solution

$$\begin{bmatrix} B \\ B \\ \vdots \\ B \\ \theta \end{bmatrix} = Q_R^{-1} Q_L \quad (\text{A.24})$$

and solutions are of the form

$$x_t = B \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix}, x_{t+1} = \frac{B}{2} \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix}, \dots, x_{t+\theta} = \frac{B}{\theta} \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix} \quad (\text{A.25})$$

If Q_R is singular, the system has an infinity of solutions.

When $\eta < L\theta$, The system has an infinity of solutions.

When Q has more than $L\theta$ rows, The system has a unique nontrivial solution only for specific values of x_{data}

Proof of rank of Q

Proof The theorem applies well known results on existence and uniqueness of solutions to linear equation systems (Noble, 1969). If $\mathcal{M}_2 = \begin{bmatrix} x_{data} \\ 0 \end{bmatrix}$ does

not lie in the column space of $\mathcal{M}_1 = \begin{bmatrix} I & 0 \\ Q_L & Q_R \end{bmatrix}$, then there is no solution. If

\mathcal{M}_2 lies in the column space of \mathcal{M}_1 and the latter matrix is full rank, then there is a unique solution. If \mathcal{M}_2 lies in the column space of \mathcal{M}_1 and the latter matrix is not full rank, then there are multiple solutions. ■

A.4 Proof of Theorem 1

Proof Construct the $L\tau \times L\tau$ matrix:

$$\tilde{B} = \begin{bmatrix} 0 & I \\ B \end{bmatrix} \quad (\text{A.26})$$

Define

$$\mathcal{B}_{k+1} = \mathcal{B}_k \tilde{B} \quad (\text{A.27})$$

Applying equation 8 to the unique convergent solution, it follows that

$$\begin{bmatrix} H_- & H_0 & H_+ \end{bmatrix} \begin{bmatrix} I \\ B \\ \vdots \\ \mathcal{B}_{\theta+1} \end{bmatrix} = 0 \quad (\text{A.28})$$

where

$$H_- = \begin{bmatrix} H_{-\tau} & \dots & H_{-1} \end{bmatrix} \quad H_+ = \begin{bmatrix} H_1 & \dots & H_\theta \end{bmatrix} \quad (\text{A.29})$$

Which can also be written as:

$$\begin{bmatrix} H_- & H_0 & H_+ \end{bmatrix} \begin{bmatrix} I \\ \begin{bmatrix} 0 & I \end{bmatrix} \\ B \\ \vdots \\ \mathcal{B}_\theta \end{bmatrix} \tilde{B} = 0 \quad (\text{A.30})$$

So that:

$$H_- + \begin{bmatrix} 0 & H_0 \end{bmatrix} + H_+ \begin{bmatrix} B \\ \vdots \\ \mathcal{B}_\theta \end{bmatrix} \tilde{B} = 0 \quad (\text{A.31})$$

$$H_- + \begin{bmatrix} 0 & H_0 \end{bmatrix} + H_+ \begin{bmatrix} B_L & B_R \\ \vdots & \vdots \\ \mathcal{B}_{L\theta} & \mathcal{B}_{R\theta} \end{bmatrix} \tilde{B} = 0 \quad (\text{A.32})$$

where the B_R^k matrices are $L \times L$.

Note that

$$\begin{bmatrix} B_L & B_R \\ \vdots & \vdots \\ \mathcal{B}_{L\theta} & \mathcal{B}_{R\theta} \end{bmatrix} \tilde{B} = \begin{bmatrix} 0 & B_L \\ \vdots & \vdots \\ 0 & \mathcal{B}_{L\theta} \end{bmatrix} + \begin{bmatrix} B_R \\ \vdots \\ \mathcal{B}_{R\theta} \end{bmatrix} B \quad (\text{A.33})$$

Now

$$H_- + H_+ \begin{bmatrix} 0 & B_L \\ \vdots & \vdots \\ 0 & \mathcal{B}_{L\theta} \end{bmatrix} + (H_0 + H_+ \begin{bmatrix} B_R \\ \vdots \\ \mathcal{B}_{R\theta} \end{bmatrix})B = 0 \quad (\text{A.34})$$

Define

$$\phi = (H_0 + H_+ \begin{bmatrix} B_R \\ \vdots \\ \mathcal{B}_{R\theta} \end{bmatrix})^{-1} \quad (\text{A.35})$$

So that

$$\phi H_- + \phi H_+ \begin{bmatrix} 0 & B_L \\ \vdots & \vdots \\ 0 & \mathcal{B}_{L\theta} \end{bmatrix} + B = 0 \quad (\text{A.36})$$

Consider the impact that the time $t + s$ value z_{t+s} has on the value of x_{t+s} . We can write

$$\begin{bmatrix} H_- & H_0 & H_+ \end{bmatrix} \begin{bmatrix} I & & & 0 \\ & \ddots & & \vdots \\ & & I & 0 \\ 0 & \dots & 0 & I \\ & & B_L & B_R \\ 0 & & \vdots & \vdots \\ & & B_L^\theta & B_R^\theta \end{bmatrix} \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix} = \Psi z_{t+s} \quad (\text{A.37})$$

or equivalently,

$$\begin{aligned} \phi \Psi z_{t+s} = & \phi \left(\begin{bmatrix} H_- & 0 \end{bmatrix} + \begin{bmatrix} 0 & H_0 \end{bmatrix} + \right. \\ & \left. \begin{bmatrix} H_+ \begin{bmatrix} 0 & B_L \\ \vdots & \vdots \\ 0 & \mathcal{B}_{L\theta} \end{bmatrix} & H_+ \begin{bmatrix} B_R \\ \vdots \\ \mathcal{B}_{R\theta} \end{bmatrix} \end{bmatrix} \right) \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix} \quad (\text{A.38}) \end{aligned}$$

or

$$\begin{aligned}
\phi\Psi z_{t+s} = & \\
& \phi\left(\begin{bmatrix} H_- & 0 \end{bmatrix} + \right. \\
& \left. \begin{bmatrix} H_+ \begin{bmatrix} 0 & B_L \\ \vdots & \vdots \\ 0 & B_{L\theta} \end{bmatrix} & 0 \end{bmatrix} \right) \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix} + \\
& \phi\left(\begin{bmatrix} 0 & H_0 \end{bmatrix} + \right. \\
& \left. \begin{bmatrix} 0 & H_+ \begin{bmatrix} B_R \\ \vdots \\ B_{R\theta} \end{bmatrix} \end{bmatrix} \right) \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix}
\end{aligned} \tag{A.39}$$

Which by equations A.35 and A.36 can be written

$$\begin{aligned}
\phi\Psi z_{t+s} = & \\
& \begin{bmatrix} -B & 0 \end{bmatrix} \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix} + \\
& \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix}
\end{aligned} \tag{A.40}$$

So we have

$$\begin{bmatrix} -B & I \end{bmatrix} \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s} \end{bmatrix} = \phi\Psi z_{t+s} \tag{A.41}$$

$$x_{t+s} = B \begin{bmatrix} x_{t+s-\tau} \\ \vdots \\ x_{t+s-1} \end{bmatrix} + \phi\Psi z_{t+s} \tag{A.42}$$

Now consider the impact of z_{t+s} on x_{t+s-1} . We can write

$$\phi\left(\begin{bmatrix} H_- & 0 \end{bmatrix} + \begin{bmatrix} 0 & H_0 \end{bmatrix} + \begin{bmatrix} H_+ \begin{bmatrix} 0 & B_L \\ \vdots & \vdots \\ 0 & \mathcal{B}_{L\theta} \end{bmatrix} H_+ \begin{bmatrix} B_R \\ \vdots \\ \mathcal{B}_\theta \end{bmatrix} \right) \begin{bmatrix} x_{t+s-\tau-1} \\ \vdots \\ x_{t+s-1} \end{bmatrix} + \phi H_+ \begin{bmatrix} I \\ B_R \\ \vdots \\ \mathcal{B}_{\theta-1} \end{bmatrix} \phi \Psi z_{t+s} = 0$$

where the last term captures the impact z_{t+s} has on values of x $t+s$ and later. Using equations A.35 and A.36 we can write

$$\begin{bmatrix} -B & I \end{bmatrix} \begin{bmatrix} x_{t+s-\tau-1} \\ \vdots \\ x_{t+s-1} \end{bmatrix} + \phi H_+ \begin{bmatrix} I \\ B_R \\ \vdots \\ \mathcal{B}_{\theta-1} \end{bmatrix} \phi \Psi z_{t+s} = 0 \quad (\text{A.43})$$

$$x_{t+s-1} = B \begin{bmatrix} x_{t+s-\tau-1} \\ \vdots \\ x_{t+s-2} \end{bmatrix} - \phi H_+ \begin{bmatrix} I \\ B_R \\ \vdots \\ \mathcal{B}_{\theta-1} \end{bmatrix} \phi \Psi z_{t+s} \quad (\text{A.44})$$

and more generally

$$x_{t+s-i} = B \begin{bmatrix} x_{t+s-\tau-1} \\ \vdots \\ x_{t+s-2} \end{bmatrix} + (-\phi H_+ \begin{bmatrix} I \\ B_R \\ \vdots \\ \mathcal{B}_{R\theta-1} \end{bmatrix})^i \phi \Psi z_{t+s} \quad (\text{A.45})$$

To accommodate lagged expectations, suppose that information on all the endogenous variables becomes available with the same lag (D^*) in time: $\exists K^*$ such that $x_{t-k} \in I_t, \forall K \geq K^*$

then set,

$$\begin{bmatrix} E[x_{t+1}|I_t] \\ \vdots \\ E[x_{t+\theta}|I_t] \end{bmatrix} = \begin{bmatrix} B \\ \vdots \\ \mathcal{B}_\theta \end{bmatrix} \tilde{B}^{k^*} \begin{bmatrix} x_{t-\tau+1-k^*}^{data} \\ \vdots \\ x_{t-k^*}^{data} \end{bmatrix} + \quad (\text{A.46})$$

$$\begin{bmatrix} \sum_{s=0}^{\infty} ((-\phi H_+ \begin{bmatrix} I \\ B_R \\ \vdots \\ \mathcal{B}_{R\theta-1} \end{bmatrix})^s \phi \Psi z_{t+s+1}) \\ \vdots \\ \sum_{s=0}^{\infty} ((-\phi H_+ \begin{bmatrix} I \\ B_R \\ \vdots \\ \mathcal{B}_{R\theta-1} \end{bmatrix})^s \phi \Psi z_{t+s+\theta}) \end{bmatrix} \quad (\text{A.47})$$

So that

$$\begin{aligned} x_t = & \begin{bmatrix} B_L & B_R \end{bmatrix} \begin{bmatrix} x_{t-\tau} \\ \vdots \\ x_{t-1} \end{bmatrix} + \\ & \begin{bmatrix} 0 & \dots & 0 & I \end{bmatrix} \sum_{s=0}^{\infty} (F^s \begin{bmatrix} 0 \\ \phi \Psi z_{t+s} \end{bmatrix}) \end{aligned}$$

Where

$$B = \begin{bmatrix} B_L & B_R \\ \vdots & \vdots \\ B_L^\theta & B_R^\theta \end{bmatrix} = Q_R^{-1} Q_L \quad (\text{A.48})$$

A.5 Proof of Theorem 2

Proof

$$\sum_{s=0}^{\infty} (F^s \begin{bmatrix} 0 \\ \phi \Psi z_{t+s} \end{bmatrix}) = \sum_{s=0}^{\infty} (F^s \begin{bmatrix} 0 \\ \phi \Psi \end{bmatrix} \Upsilon^s) z_t = \vartheta z_t \quad (\text{A.52})$$

where

$$\text{vec}(\vartheta) = \begin{bmatrix} 0 & \dots & 0 & I \end{bmatrix} (I - \Upsilon^T \otimes F)^{-1} \text{vec} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \phi \Psi \end{bmatrix} \quad (\text{A.53})$$

■

A.6 Observable Structure

Since one can write

$$\epsilon_t = \begin{bmatrix} H_{-\tau} & \dots & H_0 \end{bmatrix} \begin{bmatrix} x_{t-\tau}^{data} \\ \vdots \\ x_t^{data} \end{bmatrix} + \begin{bmatrix} H_1 & \dots & H_\theta \end{bmatrix} \begin{bmatrix} E[x_{t+1}|I_t] \\ \vdots \\ E[x_{t+\theta}|I_t] \end{bmatrix} - \Psi z_t \quad (\text{A.54})$$

We find that

$$\epsilon_t = S \begin{bmatrix} x_{t-\tau+1-\max(1,k^*)}^{data} \\ \vdots \\ x_t^{data} \end{bmatrix} - \Psi z_t \quad (\text{A.55})$$

where

$$S = \left[0_{L \times L \max(0, k^* - 1)} \ H_{-\tau} \ \dots \ H_0 \right] + \quad (\text{A.56})$$

$$\begin{bmatrix} \left[H_1 \ \dots \ H_\theta \right] \begin{bmatrix} B \\ \vdots \\ \tilde{B}^{k^*} \ 0_{L \times L \max(0, k^* - 1)} \\ \mathcal{B}_\theta \end{bmatrix} \end{bmatrix} \quad (\text{A.57})$$

Note that for $k^* \geq 1$

$$\frac{\partial \epsilon_t}{\partial x_t^{\text{data}}} = H_0 \quad (\text{A.58})$$

and for $k^* = 0$

$$\frac{\partial \epsilon_t}{\partial x_t^{\text{data}}} = H_0 + \left[H_1 \ \dots \ H_\theta \right] \begin{bmatrix} B \\ \vdots \\ \mathcal{B}_\theta \end{bmatrix} = \phi^{-1} \quad (\text{A.59})$$

A.7 Stochastic Transition Matrices

One can write

$$\begin{bmatrix} x_{t-\tau-\max(k^*-1,0)+1} \\ \vdots \\ x_t \end{bmatrix} = \mathcal{A} \begin{bmatrix} x_{t-\tau-\max(k^*-1,0)} \\ \vdots \\ x_{t-1} \end{bmatrix} + \mathcal{B} \epsilon_t \quad (\text{A.60})$$

where

$$\mathcal{A} = \begin{bmatrix} 0 & & I & & \\ \vdots & & & \ddots & \\ 0 & & & & I \\ S_0^{-1} S_{t-\tau-\max(k^*-1,0)+1} & \dots & \dots & \dots & S_0^{-1} S_{-1} \end{bmatrix} \quad (\text{A.61})$$

$$\mathcal{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_0^{-1} \end{bmatrix} \quad (\text{A.62})$$

B An Example

This section applies algorithms 1-3 to the following model with $N=2$.(Taylor, 1979)

$$w_t = \frac{1}{N} E_t \left[\sum_{i=0}^{N-1} W_{t+i} \right] - \alpha(u_t - u_n) + \nu_t \quad (\text{B.1})$$

$$W_t = \frac{1}{N} \sum_{i=0}^{N-1} w_{t-i} \quad (\text{B.2})$$

$$u_t = \vartheta u_{t-1} + \gamma W_t + \mu + \epsilon_t \quad (\text{B.3})$$

$$E_t[\nu_{t+i}] = E_t[\epsilon_{t+i}] = 0 \forall i \geq 0 \quad (\text{B.4})$$

The initial matrix $H^{\#,0}$, for the example model with variable order $\{\epsilon, \nu, u, w, W\}$ is

$$H^{\#,0} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -1 & \alpha & 1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\vartheta & 0 & 0 & -1 & 0 & 1 & 0 & -\gamma & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We would like to express the time $t + 1$ variables in terms of the time t and $t - 1$ variables. If the sub-matrix corresponding to the $t + 1$ variables were non

singular we could immediately write:

$$\begin{bmatrix} \epsilon_{t+1} \\ \nu_{t+1} \\ u_{t+1} \\ w_{t+1} \\ W_{t+1} \end{bmatrix} = (H_{\theta}^{\#,0})^{-1} \begin{bmatrix} H_{-1}^{\#,0} & H_0^{\#,0} \end{bmatrix} \begin{bmatrix} \epsilon_{t-1} \\ \nu_{t-1} \\ u_{t-1} \\ w_{t-1} \\ W_{t-1} \\ \epsilon_t \\ \nu_t \\ u_t \\ w_t \\ W_t \end{bmatrix}$$

Since $(H_{\theta}^{\#,0})$ is singular, we use equations dated subsequent to the present time period to construct a set of linear constraints where the leading block is non singular.

$$H^{\#,*} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\vartheta & 0 & 0 & -1 & 0 & 1 & 0 & -\gamma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & \alpha & 1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

So that

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2\gamma & \rho & 2\gamma & -\gamma \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 4\alpha & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 2\alpha & 2 & -1 \end{bmatrix} \quad (\text{B.5})$$

where

$$\rho = \theta + 2\gamma\alpha \quad (\text{B.6})$$

For the example model

$$Z^\sharp = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & -\vartheta & 0 & 0 & -1 & 0 & 1 & 0 & -\gamma \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & -\vartheta & 0 & 0 & -1 & 0 & 1 & 0 & -\gamma \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \phi_4 & 0 & 0 & 0 & 2 & \phi_5 & \phi_6 & 1 \end{bmatrix}$$

where the ϕ 's are

$$\begin{aligned}\phi_7 &= 1 + \gamma \phi_5 + 2 \phi_6 \\ \phi_6 &= \frac{-2 \left(18 \alpha \gamma + (-1 + \rho)^2 + (2 + \rho) \phi_2^{\frac{1}{3}} + \phi_2^{\frac{2}{3}} \right)}{\phi_3} \\ \phi_5 &= \frac{-1 - 3 \alpha \gamma + 3 \rho + 21 \alpha \gamma \rho - 3 \rho^2 + \rho^3 + \phi_1 + \frac{108 \alpha^2 \gamma^2 + (-1 + \rho)^4 + 6 \alpha \gamma (-1 + \rho)(1 + 5 \rho) + 2(-1 + \rho) \phi_1 + (12 \alpha \gamma + (-1 + \rho)^2) \phi_2^{\frac{1}{3}}}{\gamma \phi_3}}{\phi_2^{\frac{1}{3}}} \\ \phi_4 &= -\frac{3 \alpha \gamma \left(216 \alpha^2 \gamma^2 - 4(-1 + \rho)^3 + 9 \alpha \gamma (1 + (-14 + \rho) \rho) \right) + \phi_1^2}{\gamma \phi_2^{\frac{2}{3}} \phi_3} \\ \phi_3 &= 18 \alpha \gamma + (-1 + \rho)^2 + (5 + \rho) \phi_2^{\frac{1}{3}} + \phi_2^{\frac{2}{3}} \\ \phi_2 &= (-1 + \rho)^3 + 27 \alpha \gamma (1 + \rho) + 3 \phi_1 \\ \phi_1 &= \sqrt{-\left(\alpha \gamma \left(216 \alpha^2 \gamma^2 - 4(-1 + \rho)^3 + 9 \alpha \gamma (1 - 14 \rho + \rho^2) \right) \right)}\end{aligned}$$

$$B = \frac{1}{\phi_7} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \vartheta - \gamma \phi_4 + 2 \vartheta \phi_6 & \gamma \phi_6 \\ 0 & 0 & -2(\phi_4 + \vartheta \phi_5) & -1 - \gamma \phi_5 \\ 0 & 0 & -\phi_4 - \vartheta \phi_5 & \phi_6 \end{bmatrix}$$