Minimal state variable solutions to Markov-switching rational expectations models

Roger E.A. Farmer*, Daniel F. Waggoner, Tao Zha

UCLA, Federal Reserve Bank of Atlanta, Federal Reserve Bank of Atlanta, Emory University and SHUFE, United States

1. Introduction

For at least 25 years, economists have estimated structural models with constant parameters using U.S. and international data. Experience has taught us that some parameters in these models are unstable and a natural explanation for the failure of the parameter constancy assumption is that the world is changing. There are competing explanations for the source of parameter change that include abrupt breaks in the variance of structural shocks (Stock and Watson, 2003; Sims and Zha, 2006; Justiniano and Primiceri, 2008), breaks in the parameters of the private sector equations due to financial innovation (Bernanke et al., 1999; Christiano et al., 2008; Gertler and Kiyotaki, 2010), or breaks in the parameters of monetary and fiscal policy rules (Clarida et al., 2000; Lubik and Schorfheide, 2004; Davig and Leeper, 2007; Fernandez-Villaverde and Rubio-Ramirez, 2008; Christiano et al., 2009). Markov-switching rational expectations (MSRE) models can capture the fact that the structure of the economy changes over time.

Cogley and Sargent (2005a)’s estimates of random coefficient models suggest that when parameters change, they move around in a low dimensional subspace; that is, although all of the parameters of a VAR may change, they change together. This is precisely what one would expect if parameter change were due to movements in a small subset of parameters of a structural rational expectations model. Although this phenomenon can be effectively modeled as a discrete Markov process, Sims (1982) and Cooley et al. (1984) pointed out some time ago that a rational expectations model should take account of the fact that agents will act differently if they are aware of the possibility of regime change.

* This research was supported in part by NSF grants SES-0702839 and SES-1127665. We thank two referees and Michel Juillard for many thoughtful comments and Junior Maih for helpful discussions. The views expressed herein do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

* Corresponding author.
E-mail address: rfarmer@econ.ucla.edu (R.E.A. Farmer).
In a related paper (Farmer et al., 2009), we show that equilibria of MSRE models are of two types: minimal state variable (MSV) equilibria and non-fundamental equilibria. Non-fundamental equilibria may or may not exist. If a non-fundamental equilibrium exists, it is the sum of an MSV equilibrium and a secondary stochastic process. Our innovation in this paper is to develop an efficient method for finding MSV equilibria in a general class of MSRE models, including those with lagged state variables. Given the set of MSV equilibria, our earlier paper (Farmer et al., 2009) shows how to construct non-fundamental equilibria.

Previous authors, notably Leeper and Zha (2003), Svensson and Williams (2005), Davig and Leeper (2007), and Farmer et al. (2008), have made some progress in developing methods to solve for the equilibria of MSRE models. But the techniques developed to date are not capable of finding all of the equilibria in a general class of MSRE models. We illustrate this point with an example. We use a simple rational expectations model to illustrate why previous approaches (including our own) may not find an MSV equilibrium, and in the case of multiple MSV equilibria, can at best find only one MSV equilibrium. In contrast, we show that our new method is able to find all MSV equilibria. The algorithm we develop is shown to be fast and efficient.

2. Minimal state variable solutions

A general class of MSRE models studied in the literature has the following form:

\[
A_{t+1} x_t = B_{t+1} x_t + \psi x_t + \eta_t,
\]

where \(x_t\) is an \(n \times 1\) vector of endogenous and predetermined variables, \(A_t, B_t, \psi, \eta_t\), and \(x_0\) are conformable parameter matrices, \(\eta_t\) is a \(k \times 1\) vector of i.i.d. stationary exogenous shocks, and \(\eta_t\) is an \(\ell \times 1\) vector of expectation errors. The variable \(s_t\) is an exogenous stochastic process following an \(h\)-regime Markov chain, where \(s_t \in \{1, \ldots, h\}\) with transition matrix \(P = [p_{ij}]\) defined as

\[p_{ij} = P(s_t = i | s_{t-1} = j).\]

Because the vector \(\eta_t\) is a mean zero endogenous stochastic process and we implicitly assume that \(\Pi(s_t)\) is of full column rank, without loss of generality we let \(\pi_1(s_t) = 0, \pi_2(s_t) = I, \psi_1(s_t) = \psi(s_t), \) and \(\psi_2(s_t) = 0\), where \(I\) is the \(\ell \times \ell\) identity matrix.

In most applications, \(x_t\) is partitioned as

\[x_t = [y_t', z_t', \varepsilon_t', y_{t+1}'],\]

where the first pair \([y_t', z_t']\) is of dimension \(n - \ell\) and the second block of Eq. (1) is of the form \(y_t = E_{t-1} y_t + \xi_t\). In this case, the endogenous shocks \(\xi_t\) can be interpreted as expectational errors. The vector \(y_t\) is the endogenous component and \(z_t\) is the predetermined component consisting of lagged and exogenous variables. Regime-switching constant terms can be encoded by introducing a dummy variable \(z_{c,t}\) as an element of the vector \(z_t\) together with the additional equation \(z_{c,t} = z_{c,t-1}\), subject to the initial condition \(z_{c,0} = 1\). While this addition introduces a unit eigenvalue into the system, the solution techniques developed in this paper are not affected because the dummy variable is just a constant term and the stationarity of the system is intact.

In Farmer et al. (2009), we develop a set of necessary and sufficient conditions for equilibria to be determinate in a class of forward-looking MSRE models. We show in that paper that every solution of an MSRE model, including an special case of the Markov-switching system given by (1), which we represent as follows:

\[
A_{t+1} x_t = B_{t+1} x_t + \psi x_t + \eta_t,
\]

There are a variety of techniques to solve this system and the general solution is of the form

\[x_t = \Gamma x_{t-1} + E_1 \varepsilon_t + E_2 \gamma_t,\]

where the mean-zero random process \(\gamma_t\), if present, is a sunspot component. For expositional clarity, let us assume that \(A\) is invertible. The matrices \(\Gamma, E_1,\) and \(E_2\) can be obtained from the real Schur decomposition of \(A^{-1}B = \text{UTU}^T\). The matrix \(U\) is orthogonal and \(T\) is block upper triangular with \(1 \times 1\) and \(2 \times 2\) blocks along its diagonal. The \(1 \times 1\) blocks correspond to real eigenvalues of \(A^{-1}B\) and the \(2 \times 2\) blocks correspond to conjugate pairs of complex eigenvalues of \(A^{-1}B\). The real Schur
decomposition is unique up to the ordering of the eigenvalues along the block diagonal of $T$. If we partition $U$ as $U=[V \dot{V}]$, then the Schur decomposition can be written as

$$A^{-1}B = [V \dot{V}] \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} [V \dot{V}]^{-1}.$$ 

If we define $\Gamma = VT_{11}V^\prime$, $\Sigma_1 = VG_1$, and $\Sigma_2 = VN_1$, where $G_1$ and $N_1$ are solutions of the matrix equations

$$[AV II] \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = \Psi \quad \text{and} \quad [AV II] \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = 0,$$

then Eq. (4) will define a solution of the system given by (3). This is straightforward to verify by multiplying Eq. (4) by $A$ and then transforming the right hand side using the definitions of $\Gamma$, $\Sigma_1$, and $\Sigma_2$, the fact that $x_t$ is in the column space of $V$, the identity $A^{-1}BV = \Gamma V$ and the implicit definition $\eta_t = -G_2\xi_t - N_2\gamma_t$. Furthermore, any solution will correspond to some ordering of the eigenvalues $A^{-1}B$ and a partition of $U$. Since we require solutions to be stable,\(^1\) all the eigenvalues of $T_{11}$ must lie inside the unit circle.

The first requirement of an MSV solution is that it be fundamental, i.e. it cannot contain a sunspot component. This implies that $N_t$ must be zero or equivalently that $[AV II]$ must be of full column rank. The second requirement is that if $x_t$ is decomposed as an endogenous component, a predetermined component, and an accounting component as in Eq. (2), then no restrictions should be placed on the “data”, which corresponds to the endogenous and predetermined components.\(^2\) This implies that the number of columns in $V$ must be $n-\ell$ and that $[AV II]$ is invertible.

We can use these ideas to formalize what we mean by an MSV equilibrium. First, note that the column space of $V$ is the span of solution $x_t$ in the sense that support of the random process $x_t$ is contained in and spans the column space of $V$. A solution of the system (3) is an MSV solution if and only if it is the unique solution on its span and there are no restrictions on the endogenous and predetermined variables $y_t$ and $z_t$. This means that the span of the solution uniquely determines $E_2y_{t-1}$ as a function of $y_t$ and $z_t$. These ideas can be expanded to the Markov switching system given by (1) and (2). In this context, the relevant concept is not the span of the solution, but the conditional span. The span of the solution $x_t$ conditional on $s_t = i$ is the span, over all $t$, of the support of the random vector $x_t$ given $s_t = i$.

**Definition 1.** A stable solution of the system given by (1) and (2) is a minimal state variable solution if and only if it is unique given all the conditional spans and none of the conditional spans impose a relationship among the endogenous and predetermined components $y_t$ and $z_t$.

Unlike the constant parameter case, one can no longer apply an eigenvalue condition used to identify all candidates for the conditional spans. One can, however, use iterative techniques to construct MSV equilibria. Our approach builds on the following theorem.

**Theorem 1.** If the process $\{x_t, \eta_t\}_{t=1}^\infty$ is an MSV solution of the system (1), then

$$x_t = V_t F_{1,t} x_{t-1} + V_t G_{1,t} \eta_t,$$

$$\eta_t = -(F_{2,t} x_{t-1} + G_{2,t} \eta_t),$$

where the matrix $[A(i) V_t \eta_t]$ is invertible and

$$[A(i) V_t \eta_t] \begin{bmatrix} F_{1,i} \\ F_{2,i} \end{bmatrix} = B(i),$$

$$[A(i) V_t \eta_t] \begin{bmatrix} G_{1,i} \\ G_{2,i} \end{bmatrix} = \Psi(i),$$

$$(\sum_{i=1}^h p_i F_{2,i}) V_j = 0_{c,n-\ell}.$$ 

The dimension of $V_t$ is $n \times (n-\ell)$, $F_{1,i}$ is $(n-\ell) \times n$, $F_{2,i}$ is $\ell \times n$, $G_{1,i}$ is $(n-\ell) \times k$, and $G_{2,i}$ is $\ell \times k$.

Eqs. (5) and (6) define the process, Eqs. (7) and (8) ensure that the process satisfies Eq. (1), and Eq. (9) ensures that $E_{t-1}[\eta_t] = 0$. To find an MSV equilibrium, the key is to find the matrices $V_t$. With the $V_t$ in hand, Eqs. (7) and (8) can be used to find $F_{1,i}$, $F_{2,i}$, $G_{1,i}$, and $G_{2,i}$. If the $V_t$ and $F_{2,i}$ satisfy Eq. (9), then one has a candidate MSV equilibrium. It still must be verified that the solution is stationary (mean-square-stable) in the sense of Costa et al. (2004), page 36. As shown in Costa

\(^{1}\) For constant parameter systems such as (3), stable and bounded are equivalent requirements, but not so for the time varying systems such as (1).

\(^{2}\) For convenience, accounting identities may be imposed on the variables $y_t$ and $z_t$ by having a row of $e_t$ and $\psi$ equal to zero with the corresponding row of $b_t$, expressing the identity. If this is done, it is easy to see that the system is equivalent to a smaller system without the accounting identity. Thus we assume that $a_t$ or $a_t(i)$ is of full row rank.
et al. (2004), Proposition 3.9, p. 36 and Proposition 3.33, p. 49, the candidate MSV solution is stationary if and only if the eigenvalues of

\[(P \otimes I_n) \text{diag}(V_1 F_{1,1} \otimes V_1 F_{1,1}, \ldots, V_b F_{1,b} \otimes V_b F_{1,b})\]  

(10)

are all inside the unit circle.\(^3\)

Since \( \Pi = [0 I_{n-\ell}, I_\ell] \), the matrix \( [A(i)V_i] \Pi \) is invertible if and only if the upper \((n-\ell) \times (n-\ell)\) block of \( A(i) V_i \) is invertible. It is easy to see that multiplying \( V_i \) on the right by an invertible matrix, and hence multiplying \( F_{1,i} \) and \( G_{1,i} \) on the left by the inverse of this matrix, will not change Eq. (5) through (9). Thus, without loss of generality, we assume that

\[A(i) V_i = \begin{bmatrix} I_{n-\ell} \\ -X_i \end{bmatrix}\]  

(11)

for some \((n-\ell) \times (n-\ell)\) matrix \(X_i\). Since

\[F_{2,i} = [0 I_{n-\ell}, I_\ell][A(i)V_i] \Pi^{-1} B(i) = [X_i I_\ell] B(i),\]

Eq. (9) becomes

\[
\sum_{i=1}^{b} p_{ij} [X_i I_\ell] B(i) A(j)^{-1} \begin{bmatrix} I_{n-\ell} \\ -X_j \end{bmatrix} = 0_{n-\ell, n-\ell}.
\]

(12)

In this derivation, we have assumed that \( A(i) \) is invertible for expositional clarity. In Appendix B, we remove this assumption and show that our iterative algorithm works even if \( A(i) \) is not invertible.

The advantage of our method is that we are able to reduce the task of finding an MSV solution to that of computing the roots of a quadratic polynomial in several variables. We exploit Newton’s method to compute these roots. This has the advantage over previously suggested methods of being fast and locally stable around any given solution. This property guarantees that by choosing a large enough grid of initial conditions we will find all possible MSV solutions. This local convergence property does not hold for iterative solutions that have previously been suggested in the literature.

Let \( X = (X_1, \ldots, X_b) \), define \( f_j \) to be the function from \( \mathbb{R}^{b(n-\ell)} \) to \( \mathbb{R}^{n(n-\ell)} \) given by

\[f_j(X) = \sum_{i=1}^{b} p_{ij} [X_i I_\ell] B(i) A(j)^{-1} \begin{bmatrix} I_{n-\ell} \\ -X_j \end{bmatrix},\]

(13)

and \( f \) be the function from \( \mathbb{R}^{b(n-\ell)} \) to \( \mathbb{R}^{n(n-\ell)} \) given by

\[f(X) = (f_1(X), \ldots, f_b(X)).\]

(14)

The quadratic polynomial equations, \( f(X) = 0 \), are the same as the constraints represented by (9).

Thus, finding an MSV equilibrium is equivalent to finding the roots of \( f(X) \) and Theorem 1 suggests the following constructive algorithm for finding MSV solutions.

**Algorithm 1.** Let \( X^{(k)} = (X_1^{(k)}, \ldots, X_b^{(k)}) \) be an initial guess. If the \( k \)th iteration is \( X^{(k)} = (X_1^{(k)}, \ldots, X_b^{(k)}) \), then the \((k+1)\)th iteration is given by

\[\text{vec}(X^{(k+1)}) = \text{vec}(X^{(k)}) - f'(X^{(k)})^{-1} \text{vec}(f(X^{(k)})�)\]

where

\[f'(X) = \begin{bmatrix}
\frac{\partial f_1(X)}{\partial X_1} & \ldots & \frac{\partial f_1(X)}{\partial X_b} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_b(X)}{\partial X_1} & \ldots & \frac{\partial f_b(X)}{\partial X_b}
\end{bmatrix}.
\]

The sequence \( X^{(k)} \) converges to a root of \( f(X) \).

It is straightforward to verify that for \( i \neq j \)

\[\frac{\partial f_j}{\partial X_i}(X) = p_{ij} \left( [I_{n-\ell} 0_{n-\ell,\ell}] B(j) A(j)^{-1} \begin{bmatrix} I_{n-\ell} \\ -X_j \end{bmatrix} \right)^\prime \otimes I_{\ell}\]

and for \( i = j \),

\[\frac{\partial f_j}{\partial X_j}(X) = p_{jj} \left( [I_{n-\ell} 0_{n-\ell,\ell}] B(j) A(j)^{-1} \begin{bmatrix} I_{n-\ell} \\ -X_j \end{bmatrix} \right)^\prime \otimes I_{\ell} + \sum_{k=1}^{b} p_{kj} [X_k I_\ell] B(k) A(j)^{-1} \begin{bmatrix} 0_{n-\ell,\ell} \\ -I_{\ell} \end{bmatrix}.\]

\(^3\) If there is a constant term, then there will be roots on the unit circle.
In a series of computational experiments, reported below, we have found that this algorithm is relatively fast and that it converges to multiple solutions, when they exist, for a suitable choice of initial conditions.

In Section 4, we present simple examples in which existing algorithms, that have been proposed in the literature, break down. We also show that when there are multiple MSV equilibria, existing algorithms can at best find only one equilibrium and sometimes do not converge to any MSV equilibrium even when the initial starting point is close to the equilibrium. This result is unsatisfactory because researchers should be able to estimate models by searching across the space of all equilibria and selecting the one that maximizes the posterior odds ratios. In all the examples we study, our algorithm is capable of finding all MSV equilibria by randomly choosing different initial points.

3. Previous approaches

Two existing algorithms have been frequently used to find an MSV equilibrium in a MSRE model: the fixed-point (FP) algorithm developed in a previous version of this paper (Farmer et al., 2008) and the iterative algorithm proposed by Svensson and Williams (2005). We review these algorithms in this section and in Section 4 we discuss why they do not always work well in practice.

3.1. The FP algorithm

To apply the FP algorithm, Farmer et al. (2008) show how to define an expanded state vector \( \tilde{x}_t \). Using their definition, one can write the Markov switching equations as a constant parameter system of the form

\[
\tilde{A}\tilde{x}_t = \tilde{B}\tilde{x}_{t-1} + \tilde{\Psi}\tilde{u}_t + \tilde{H}\eta_t,
\]

where \( \tilde{x}_t \in \mathbb{R}^{nh} \) has dimension \( nh \times 1 \).

To write system 1 in this form, define a family of matrices \( \phi_i \) where \( h \) is the number of Markov states and each \( \phi_i \) is a solution of the original nonlinear system. Define \( e_i \) as a column vector equal to 1 in the \( i \)th element and zero everywhere else and the matrix \( \Phi \) as

\[
\Phi = \left[ \begin{array}{c} e_1 \otimes \phi_1 \\ \vdots \\ e_h \otimes \phi_h \end{array} \right].
\]

Let the matrices \( \tilde{A}, \tilde{B}, \) and \( \tilde{H} \) be given by

\[
\tilde{A}_{nh \times nh} = \begin{bmatrix} \text{diag}(a_1(1), \ldots, a_1(h)) \\ \vdots \\ a_2 \cdots a_2 \end{bmatrix},
\]

\[
\tilde{B}_{nh \times nh} = \begin{bmatrix} \text{diag}(b_1(1), \ldots, b_1(h))(P \otimes I_n) \\ b_2 \cdots b_2 \end{bmatrix},
\]

\[
\tilde{H}_{nh \times \ell} = [0, I_\ell, 0]'.
\]

To define \( \tilde{u}_t \) and the corresponding coefficient matrix \( \tilde{\Psi} \), let \( I_n \) be the \( h \)-dimensional column vector of ones and let

\[
S_{(n-\ell)\times nh} = (\text{diag}(b_1(1), \ldots, b_1(h))) \times [(e_i I_n - P) \otimes I_n],
\]

for \( i = 1, \ldots, h \). With this notation, we have

\[
\tilde{u}_t = \begin{bmatrix} S_{(n-\ell)\times nh} \left( e_{(n-\ell)i} \otimes (I_n \otimes I_\ell) \right) \\ e_i \otimes u_t \end{bmatrix}
\]

and

\[
\tilde{\Psi}_{nh \times (k + n - \ell)h} = \begin{bmatrix} I_{(n-\ell)h} \text{diag}(\psi(1), \ldots, \psi(h)) \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.
\]

It is straightforward to show that \( E_{t-1}[u_t] = 0 \). Thus, (15) is a linear system of rational expectations equations and the solution of this linear system can be computed by known methods. Farmer et al. (2008) show that a solution of the expanded system (15) with the initial conditions \( x_0 \) and \( \tilde{x}_0 = e_i \otimes x_0 \) is a solution of the original nonlinear system. The vectors \( x_t \) and \( \tilde{x}_t \) are related by the expression:

\[
x_t = (e_i \otimes I_n)\tilde{x}_t.
\]
Although (3) is a linear rational expectations system, finding \((\phi_1, \phi_2, \ldots, \phi_n)\) for this linear system is a fixed-point problem of a system of nonlinear equations. Farmer et al. (2008) propose the following algorithm. Let the superscript \(n\) denote the \(n\)th step of an iterative procedure. Beginning with a set of initial matrices \(\{\phi_{j,n}\}_{j=1}^{2}\), define \(\Phi^{(n)}\) using Eq. (16) and generate the associated matrix \(A^{(n)}\). Next, compute the QZ decomposition of \((A^{(n)}, B)\) and denote the generalized eigenvalues corresponding the unstable roots by \(Z_{i}^{(n)} = [z_{i1}^{(n)}, \ldots, z_{in}^{(n)}]\), where \(z_{i}^{(n)}\) is an \(\ell \times n\) matrix. Finally, set \(\phi_{j}^{(n+1)} = z_{j}^{(n)}\). Form this new set of values of \(\phi_{j}\)'s, form a new matrix \(A^{(n+1)}\). Repeat this algorithm and, if it converges, the system (15) will generate sequences \((x_t, y_t)_{t=1}^{\infty}\) that are consistent with the system (1), where \(x_t\) is governed by (17).

The qualification if it converges is crucial because, as we will show in Section 4, it may not converge even in the simplest rational expectations model.

3.2. The SW algorithm

In this subsection we describe the algorithm developed by Svensson and Williams (2005). As we exhaust many commonly used mathematical symbols for matrices and vectors, we will use the same notation for some variables and parameters as in Section 3.1 as long as this double use of the notation does not cause confusion.

Svensson and Williams (2005)'s algorithm is an iterative approach to solving a general Markov-switching system. The system is written as

\[
X_t = A_{11,t} X_{t-1} + A_{12,t} x_{t-1} + C_n e_t,
\]

(18)

\[
E_t H_{s_t} x_{t+1} = A_{21,s_t} X_t + A_{22,s_t} x_t,
\]

(19)

where \(X_t\) is an \(n_x \times 1\) vector of predetermined variables, \(x_t\) is an \(n_x \times 1\) vector of forward-looking variables, and \(s_t\). The MSV solution takes the following form:

\[
x_t = G_n X_t.
\]

The algorithm works as follows:

1. Start with an initial guess of \(G_{j}^{(0)}\), where \(s_j = j\).
2. For \(n = 0, 1, 2, \ldots\), iterate the value of \(G_{j}^{(n+1)}\) according to

\[
G_{j}^{(n+1)} = \left[ A_{22,j} - \sum_k P_{kj} H_k G_k^{(n)} A_{12,k} \right]^{-1} \left[ \sum_k P_{kj} H_k G_k^{(n)} A_{11,k} - A_{21,j} \right].
\]

(20)

This algorithm is both elegant and efficient and can handle a large system. If it converges to an MSV solution, the convergence is fast. As we show below, however, the algorithm may not converge even if there is an MSV equilibrium.

4. Comparison of our algorithm with alternatives

In this section we illustrate the properties of different methods using three simple examples based on the following model:

\[
\phi_n \pi_t = E_t \pi_{t+1} + \delta_n \pi_{t-1} + \beta_n r_t,
\]

\[
r_t = \rho_n r_{t-1} + e_t,
\]

where \(s_t = 1, 2\) takes one of two discrete values according to the Markov-switching process. If we interpret \(\pi_t\) as inflation and \(r_t\) as an exogenous shock to income or preferences, this equation can be derived directly from the consumer’s optimization problem together with a monetary policy rule that moves the interest rate in response to current and past inflation rates (see Liu et al., 2009).

4.1. An example with a unique MSV equilibrium

We set \(\delta_n = 0, \beta_n = \beta = 1, \) and \(\rho_n = \rho = 0.9\) for all values of \(s_n\). \(\phi_1 = 0.5, \phi_2 = 0.8, p_{11} = 0.8, \) and \(p_{22} = 0.9\). One can show that for this parameterization (i.e., \(\delta_n = 0\)), there is a unique MSV equilibrium. The MSV solution has a closed form given by the expression

\[
\pi_t = g_{1,s_t} r_{t-1} + g_{2,s_t} e_t,
\]

where

\[4\] There also exists a continuum of non-fundamental equilibria around the unique MSV solution.
where

\[
\begin{bmatrix}
G_{1,1} \\
G_{1,2}
\end{bmatrix} = - \begin{bmatrix}
p_{11}\rho - \phi_1 & p_{21}\rho \\
p_{12}\rho & p_{22}\rho - \phi_2
\end{bmatrix}^{-1} \begin{bmatrix}
\beta \rho \\
\beta \rho
\end{bmatrix},
\]

\[
G_{2,n} = \frac{p_{1s}G_{1,1} + p_{2s}G_{1,2} + \beta}{\phi_{s,1}}.
\]

In experiments based on this example, our algorithm converged quickly to the following MSV equilibrium for all initial conditions:

\[
\pi_t = -10.92857 + 12.14286 \epsilon_t \quad \text{for } s_t = 1,
\]

\[
\pi_t = 8.35714 + 9.28571 \epsilon_t \quad \text{for } s_t = 2.
\]

Using (10), one can easily verify that this equilibrium is mean square stable.

Both the FP or the SW algorithms, however, are unstable when applied to this example. To gain an intuition of why these previous algorithms do not work, we map this example to the notation of the SW algorithm described in Section 3.2:

\[
H_k = 1, \quad n_X = n_s = 1, \quad X_t = r_t, \quad x_t = \pi_t, \quad A_{11,k} = \rho, \quad A_{12,k} = 0, \quad A_{21,j} = -\beta, \quad A_{22,j} = \phi_j.
\]

For expositional clarity, we further simplify the model by assuming that \(\phi_1 = \phi_2 = \phi = 0.85\). The MSV equilibrium for this case can be characterized as

\[
\pi_t = g_1 r_{t-1} + g_2 \epsilon_t,
\]

where \(g_1 = \beta \rho / (\phi - \rho)\). It follows from (20) that

\[
g_1^{(n)} = \left(\frac{g_1^{(n-1)} + \beta \rho}{\phi}\right).\]

The above iterative algorithm also characterizes the FP algorithm. Since the MSV solution \(g_1\) is greater than 1 in absolute value and \(\rho / \phi > 1\) in this case, \(g_1^{(n)}\) will go to either plus infinity or minus infinity (depending on the initial guess) as \(n \to \infty\). Thus, the FP and SW algorithms cannot find the MSV equilibrium, even when there is only a unique MSV equilibrium.

4.2. An example with two MSV equilibria

We now provide an example where there are multiple MSV equilibria, but the SW algorithm can find only one of the two MSV equilibria and the FP algorithm cannot converge at all. In contrast, our proposed algorithm converges to all of the MSV equilibria by randomly selecting different sets of initial guesses. The example has the following parameter configuration:

\[
\phi_1 = 0.5, \quad \phi_2 = 0.8, \quad \delta_1 = -0.7, \quad \delta_2 = 0.4,
\]

\[
\beta_1 = \beta_2 = 1, \quad \rho_1 = \rho_2 = 0, \quad p_{11} = 1.0, \quad p_{22} = 0.64.
\]

One can easily verify that the first regime, taken in isolation, is determinate while the second regime is indeterminate. We choose this example to show that even though the first regime is an absorbing state because \(p_{11} = 1.0\), the MSV equilibrium in the regime-switching environment is not unique. To see this point clearly, note that the MSV solution takes the form \(\pi_t = g_{1n} \pi_{t-1} + g_{2n} \epsilon_t\) with two distinct stationary equilibria:

\[
g_{1,1} = -0.623212, \quad g_{1,2} = 0.675998, \quad \text{first MSV equilibrium};
\]

\[
g_{1,1} = -0.623212, \quad g_{1,2} = 0.924559, \quad \text{second MSV equilibrium}.
\]

Note that the multiple equilibria occur only in the second regime. The equilibrium in the first regime is unique.

The SW algorithm cannot find the second equilibrium; it converges only to the first equilibrium. The FP algorithm fares worse. It cannot converge to either of the two MSV equilibria.

4.3. An example with more than two MSV equilibria

We now provide an example that a multiplicity of MSV equilibria can exist. Both FP and SW algorithms can find only one of them. The question is whether our proposed algorithm is capable of finding all the solutions or only a subset of them.

The example has the following parameter configuration:

\[
\phi_1 = 0.2, \quad \phi_2 = 0.4, \quad \delta_1 = -0.7, \quad \delta_2 = -0.2,
\]

\[
\beta_1 = \beta_2 = 1, \quad \rho_1 = \rho_2 = 0, \quad p_{11} = 0.9, \quad p_{22} = 0.8.
\]
An MSV equilibrium takes the form \( p_t = g_1, s_t p_{t-1} + g_{2, s_t} e_t \). For this example, there are four stationary MSV equilibria given by
\[
\begin{align*}
g_{1,1} &= -0.765149, & g_{1,2} &= -0.262196, & \text{first MSV equilibrium}; \\
g_{1,1} &= 0.960307, & g_{1,2} &= 0.646576, & \text{second MSV equilibrium}; \\
g_{1,1} &= -0.826316, & g_{1,2} &= 0.96551, & \text{third MSV equilibrium}; \\
g_{1,1} &= 1.024809, & g_{1,2} &= -0.392746, & \text{fourth MSV equilibrium}.
\end{align*}
\]

Our algorithm converges rapidly to all the MSV solutions when we vary the initial guess randomly. In contrast, both the FP and SW algorithms, no matter what the initial guess (unless it is set exactly at an MSV solution), converge to only the first MSV equilibrium reported above.

Farmer et al. (2008) show an easy-to-check condition for the uniqueness of the equilibrium if it is found by the FP algorithm. This condition applies only to local uniqueness and to the stacked linear system (15). This local result cannot be extended to the original Markov-switching system (1). Indeed, as this example shows, even the first MSV equilibrium is locally unique according to Farmer et al. (2008), there exist other MSV equilibria that are not in the neighborhood of the first equilibrium. Our new method is developed to find all possible MSV equilibria.

5. A general strategy of selecting an equilibrium

In this section we discuss a general strategy of selecting an equilibrium in the presence of multiple MSV equilibria. We first provide details of a new efficient algorithm that we use to draw initial guesses that cover a wide range of values in order to find all the MSV equilibria. After we have all the MSV equilibria in hand, we then propose a likelihood based criterion for selecting an MSV equilibrium while discussing other alternative criteria.

5.1. Initial values

Our algorithm requires an initial guess. A brute force approach is to simply use a large grid of initial values in a hope that different initial values may lead to different MSV equilibria. This approach is not a problem for a theoretical paper whose purpose is to highlight key properties of a particular model of interest. In an estimation exercise, however, this approach can become extremely inefficient when the size of a dynamic stochastic general equilibrium (DSGE) model is large.

An efficient approach is to randomly sample initial values by exploring the theoretical properties of the MSV solution. From the solution (5) one can see that \( V_i \) is uniquely determined only up to a normalization discussed in Hamilton et al. (2007) for cointegrated systems. Thus, we can always impose the restriction that the columns of \( V_i \) be orthonormal, even though we used a different normalization in our iterative technique for finding solutions. Theorem 9 in Rubio-Ramirez et al. (2010) gives an efficient algorithm for implementing a uniform random selection of \( V_i \). Specifically, let \( \tilde{X}_i \) be an \( n \times n \) random matrix with each element having an independent standard normal distribution; and let \( \tilde{X}_i = Q_i \tilde{R}_i \) be the QR decomposition of \( \tilde{X}_i \) with the diagonal of \( \tilde{R}_i \) normalized to be positive. Then the first \( n-\ell \) columns of \( Q_i \) form an independent uniform random selection of \( V_i \). The following algorithm gives a systematic way of finding all MSV equilibria.

**Algorithm 2.** For each independent selection of \( V_i \), obtain the corresponding random selection of the initial value of \( X_i \) by multiplying by the inverse of the upper \( n-\ell \times n-\ell \) block of \( A(i)V_i \).

(Step 1) Randomly draw \( \hat{N} \) initial values of \( (X_1, \ldots, X_{\hat{N}}) \).
(Step 2) For each initial value, apply Algorithm 1 to find an MSV equilibrium.
(Step 3) Collect all MSV equilibria.
(Step 4) Repeat Steps 1–3 with \( \hat{N} = 2n\hat{N} \) initial values.
(Step 5) Compare all MSV equilibria in Step 4 to the previously obtained MSV equilibria.
(Step 6) If they are the same, stop. If there are additional MSV equilibria, go back to Steps 4 and 5.

Our experience indicates that with the starting number \( \hat{N} = 20 \), it often takes no more than three repetitions for Algorithm 2 to converge.

5.2. How to select a particular MSV equilibrium?

Once we obtain all MSV equilibria, a relevant question is: Which equilibrium should be selected? One answer is to follow the engineering literature (Costa et al., 2004) and select the MSV equilibrium that is most stationary (i.e., the equilibrium with the smallest dominant eigenvalue (in absolute value) of the matrix (10)). The intuition is that this most stationary is likely to be most “attractive” in the sense that most initial guesses of \( X \) will converge to this equilibrium.
It turns out that this intuition is not always true. To see this point, we conduct a heuristic exercise by randomly selecting 1000 initial values of X and tabulating the percentage in which a particular equilibrium the initial values converge to. For the example discussed in Section 4.2, the first equilibrium (with the dominant eigenvalue 0.388) receives 73% and the second equilibrium (with the dominant eigenvalue 0.547) receives 27%. For the example studied in Section 4.3, the first and second equilibria (with the dominant eigenvalues being 0.529 and 0.845 respectively) share the highest percentage of convergence and each receives 33%. The second highest percentage of convergence, 26%, goes to the third equilibrium (with the dominant eigenvalue 0.811). The fourth equilibrium (with the dominant eigenvalue 0.949) has the lowest percentage of convergence (8%). This example shows that a less stationary equilibrium can have the highest degree of attraction.

A better argument for selecting the most stationary MSV equilibrium is offered by Ellison and Pearlman (2010). They show that the most stationary MSV equilibrium is E-stable while other equilibria are not. This is a persuasive argument from the view point of learning. For Markov-switching rational expectations models themselves, however, a more relevant question is based on the likelihood principle: Which equilibrium should be selected conditional on the data we observe? This alternative question is important because, ultimately, an equilibrium we select ought to explain the observed data.

We propose the following likelihood based approach. For each configuration of model parameters, we use Algorithms 1 and 2 to find all MSV equilibria. For each equilibrium, we compute the likelihood value recursively by following the method of Sims et al. (2008) (note that the prior density value is the same for all the equilibria). We compare all the likelihood values and select an equilibrium associated with the highest likelihood value. It is important to bear in mind that for a different configuration of model parameters due to parameter uncertainty, the nature of the selected equilibrium may be different as well.

6. An application to a monetary policy model

In previous sections, we showed that the FP and SW algorithms may not converge to an MSV equilibrium and that if they converge, they converge to only one MSV equilibrium. In contrast, our new algorithm, using Newton’s method to compute roots, is stable, efficient, and reliable for finding all MSV equilibria.

In this section we present simulation results based on a calibrated version of the New-Keynesian model and we use it to study changes in output, inflation, and the nominal interest rate. Clarida et al. (2000) and Lubik and Schorfheide (2004) argue that the large fluctuations in output, inflation, and interest rates are manifestations of indeterminacy induced by passive monetary policy. Sims and Zha (2006), on the other hand, find no evidence in favor of indeterminacy when they allow monetary policy to switch regimes stochastically. Furthermore, they find that once the model permits time variation in disturbance variances, there is no evidence in favor of policy changes at all (see also Cogley and Sargent, 2005b; Primiceri, 2005).

Once it is known that policy changes might occur, a rational agent should treat these changes probabilistically and the probability of a future policy change should enter into his current decisions. Previous work in this area has neglected these effects and all of the studies cited above study regime switches in a purely reduced form model. We show in this section how to use the MSV solution to a MSRE model to study the effects of regime change that is rationally anticipated to occur. We use simulation results to show that the persistence and volatility in inflation and the interest rate can be the result of policy changes at all (see also Cogley and Sargent, 2005b; Primiceri, 2005).

Our regime-switching policy model, based on Lubik and Schorfheide (2004), has the following three structural equations:

\[
X_t = E\tau X_{t+1} - \tau S_t (R_t - E\tau T_{t+1}) + Z_{D,t},
\]

(21)

\[
\pi_t = \beta(S_t)E\pi_{t+1} + \kappa(S_t)X_t + Z_{S,t},
\]

(22)

\[
R_t = \rho(S_t)R_{t-1} + (1-\rho(S_t))\gamma_1(S_t)\pi_t + \gamma_2(S_t)X_t + \epsilon_{R,t},
\]

(23)

where \(X_t\) is the output gap at time \(t\), \(\pi_t\) is the inflation rate, and \(R_t\) is the nominal interest rate. Both \(\pi_t\) and \(R_t\) are measured in terms of deviations from the steady state. The coefficient \(\tau\) measures the intertemporal elasticity of substitution, \(\beta\) is the household’s discount factor, and the parameter \(\kappa\) reflects the rigidity or stickiness of prices.

The shocks to the consumer and firm’s sectors, \(Z_{D,t}\) and \(Z_{S,t}\), are assumed to evolve according to an AR(1) process:

\[
\begin{bmatrix}
Z_{D,t} \\
Z_{S,t}
\end{bmatrix} =
\begin{bmatrix}
\rho_D(S_t) & 0 \\
0 & \rho_S(S_t)
\end{bmatrix}
\begin{bmatrix}
Z_{D,t-1} \\
Z_{S,t-1}
\end{bmatrix} +
\begin{bmatrix}
\epsilon_{D,t} \\
\epsilon_{S,t}
\end{bmatrix},
\]

where \(\epsilon_{D,t}\) is the innovation to a demand shock, \(\epsilon_{S,t}\) is an innovation to the supply shock, and \(\epsilon_{R,t}\) is a disturbance to the policy rule. All these structural shocks are i.i.d. and independent of one another. The standard deviations for these shocks are \(\sigma_D(S_t)\), \(\sigma_S(S_t)\), and \(\sigma_R(S_t)\).

5 Their theoretical results pertain only to a class of rational expectations models without Markov-switching parameters.

6 See Liu et al. (2009) for a proof that the steady state in this example does not depend on regimes.
Lubik and Schorfheide (2004) estimate a constant-parameter version of this model for the two subsamples: 1960:I–1979:II and 1979:III–1997:IV. In our calibration we consider two regimes. The parameters in the first regime correspond to their estimates for the period 1960:I–1979:II and the parameters in the second regime correspond to those for 1979:III–1997:IV. The calibrated values are reported in Tables 1 and 2. The transition matrix is calculated by matching the average duration of the first regime to the length of the first subsample and by assuming that the second regime is absorbing to accommodate the belief that the pre-Volcker regime will never return7:

\[
P = \begin{bmatrix}
0.9872 & 0 \\
0.0128 & 1
\end{bmatrix}.
\]

A simple calculation verifies that, if only one regime were allowed to exist (in the sense that a rational agent was certain that no other policy would ever be followed) the first regime would be indeterminate and the second would be determinate. When a rational agent forms expectations by taking account of regime changes, we need to know if there exist multiple MSV equilibria. In our computations we apply our method to this system with a large number of randomly selected starting points and we obtain multiple MSV solutions for some configurations of parameterization that we report below.

This kind of forward-looking model provides a natural laboratory to experiment with different scenarios in light of the debate on changes in policy or changes in shock variances. The estimates provided by Lubik and Schorfheide (2004) and reported in Tables 1 and 2 mix changes in coefficients related to monetary policy with changes in other parameters in the model, since Lubik and Schorfheide (2004) do not account for the effect of the probability of regime change on the current behavior. One variation in the structural parameter values is to let the coefficient on the inflation variable in the policy Eq. (23) change while holding all the other parameters fixed across the two regimes. Tables 3 and 4 report the parameter values corresponding to this scenario, in which all the other parameters take the average of the values in Tables 1 and 2 over the two regimes. We call this scenario “policy change only”.

In a second scenario, “variance change only”, we keep the value of the policy coefficient \(\gamma_1\) at 2.19 for both regimes while letting the standard deviation \(\sigma_D\) in the first regime be five times larger than that in the second regime and keeping the value of \(\sigma_S\) at 0.3712 for both regimes.\(^8\) The parameter values for this scenario are reported in Tables 5 and 6.

---

7 One could also match the average duration of the second regime to the length of the second subsample, which give \(p_{12} = 0.9865\).

8 Sims and Zha (2006) find that differences in the shock standard deviation across regimes can be on the scale of as high as 10–12 times. One could also decrease the difference in \(\sigma_D\), and increase the difference in \(\sigma_S\) or experiment with different combinations. Our result that changes in variances matter a great deal will hold.
The last scenario we consider allows only the parameters in the private sector to change. We call it “private-sector change only”. The idea is to study whether the persistence and volatility in inflation can be generated by the changes in the private sector in a forward-looking model. We let the coefficient $t$ be 0.06137 in the first regime and 0.6137 in the second regime. Tables 7 and 8 report the values of all the parameters for this scenario. Similar results can be achieved if one lets the value of $k$ in the first regime be much smaller than that in the second regime.

Using the method discussed in Section 2, we obtain two MSV equilibria that characterize the first two scenarios and a unique MSV equilibrium for the last two scenarios. Figs. 1–3 display simulated paths of the output gap, the interest rate, and inflation under each of these scenarios. With the original estimates reported in Lubik and Schorfheide (2004), the largest eigenvalue for the matrix (10) is 0.8617 for one equilibrium and 0.7225 for the other. The dynamics are quite different for these two MSV equilibria. We display the simulated data based on the MSV equilibrium with the largest eigenvalue 0.8617. The top chart in Fig. 1 shows that the output gaps in the first regime display persistent and large
fluctuations relative to their paths in the second regime. It is well known that the constant-parameter New-Keynesian model of this type is incapable of generating much of the difference in output volatility between the two regimes. This is certainly true for the equilibrium with the largest eigenvalue 0.7225. When taking regime switching into account, we have two MSV equilibria and the difference in output dynamics between two regimes shows up in one of the equilibria.

When we restrict changes to the policy coefficient $g_1$ only, the results are very similar to the first scenario, implying it is the change in policy across regimes that causes macroeconomic dynamics to be different across regimes. For this policy-change-only scenario, we have two MSV equilibria, one with the largest eigenvalue of the matrix (10) being 0.8947 and the other equilibrium with 0.6972. The second chart from the top in Fig. 1 report the dynamics of output in the MSV equilibrium with the largest eigenvalue 0.6972. As one can see, the volatility in output is similar across the two regimes. In summary, the top two charts in Fig. 1 demonstrate that one can obtain rich dynamics from different MSV equilibria. Thus, it is important that a method be capable of finding all MSV equilibria if one would like to confront the model with the data.

When we allow only variances to change (the third scenario), there is a unique MSV equilibrium. The solution to this model is obtained by using the standard solution method of Sims (2002) because $E_{t-1}e_{1t} = 0$ for $i \in \{R,D,S\}$ even though their variances switch regime and because the uniqueness of a solution depends only on the parameters that are time invariant. As one can see from the third chart in Fig. 1, the volatility of output in the first regime is distinctly larger than that in the second regime. The difference in volatility of output across regimes disappears in the private-sector-change-only scenario (the fourth scenario), as shown in the bottom chart of Fig. 1.

Fig. 1. Simulated output gap paths from our regime-switching forward looking model. The shaded area represents the first regime.
Figs. 2–3 display the simulated dynamics of the interest rate and inflation for the four scenarios. In all scenarios, both inflation and the interest rate in the first regime display persistent and large fluctuations relative to their paths in the second regime. The degree of persistence and volatility in these variables in the first regime increases with persistence of the shock $z_{D,t}$ or $z_{S,t}$ and with the size of shock variance $\sigma_{D,t}$ or $\sigma_{S,t}$. Our final scenario is particularly interesting because, as illustrated by the bottom charts of Figs. 2–3, even if there is no change in policy and in shock variances, inflation and the interest rate can have much larger fluctuations in the first regime than in the second regime when the parameters of the private sector equations are allowed to change across regimes.

These examples teach us that the sharply different dynamics in output, the interest rate, and inflation observed before and after 1980 could potentially be attributed to different sources. The methods we have developed here give researchers the tools to address this and other issues in a regime-switching rational expectations in which rational agents take into account the probability of regime change when forming their expectations.

7. Conclusion

We have developed a new approach to solving a general class of MSRE models. The algorithm we have developed has proven efficient and reliable in comparison to the previous methods. We have shown that MSV equilibria can be characterized as a vector-autoregression with regime switching, of the kind studied by Hamilton (1989) and Sims and Zha (2006). Our new method provides tools necessary for researchers to solve and estimate a variety of regime-switching DSGE models.
Appendix A. Proof of theorem 1

Let \( (x_t, \eta_t)_{t=1}^{\infty} \) be any solution of Eq. (1). Let \( V_i \) be any \( n \times k_i \) matrix whose columns form a basis for the span of this solution conditional on \( s_t = i \). Applying the \( E_{t-1}[\cdot|s_t = i] \) operator to Eq. (1) gives

\[
A(i)E_{t-1}[x_t|s_t = i] = B(i)x_{t-1} + PTE_{t-1}[\eta_t|s_t = i].
\] (A.1)

Because \( E_{t-1}[x_t|s_t = i] \) is in the column space of \( V_i \) and the span of \( x_{t-1} \) conditional on \( s_{t-1} = j \) is the column space of \( V_j \), there exist \( k_i \times k_j \) matrices \( F_{1,i,j} \) and \( \ell \times k_j \) matrices \( F_{2,i,j} \) such that

\[
[A(i)V_i P] \begin{bmatrix} F_{1,i,j} \\ F_{2,i,j} \end{bmatrix} = B(i)V_j.
\] (A.2)

Furthermore, since

\[
\sum_{i=1}^{h} p_{i,i-1} E_{t-1}[\eta_t|s_t = i] = E_{t-1}[\eta_t] = 0,
\]

Fig. 3. Simulated inflation paths from our regime-switching forward looking model. The shaded area represents the first regime.
we can choose the $F_{1,i,j}$ and $F_{2,i,j}$ so that
\[ \sum_{i=1}^{h} p_{ij} F_{2,i,j} = 0 \cdot k_i. \]  
(A.3)

Subtracting Eq. (A.1) from Eq. (1) gives
\[ A(i)x_t - E_{L-1} x_t | S_t = i) = \Psi(i)x_t + \Pi(\tau_{1-L} E^*_{L-1} | S_t = i). \]

This implies that there exist $k \times k$ matrices $G_{1,i,j}$ and $\ell \times k$ matrices $G_{2,i,j}$ such that
\[ [A(i)V, \Pi] \begin{bmatrix} G_{1,i,j} \\ G_{2,i,j} \end{bmatrix} = \Psi(i). \]  
(A.4)

Let $V^*$ denote the generalized inverse of $V$. Because $x_t$, conditional on $s_t = i$, is in the column space of $V_s$, there exists a $k_s$ dimensional vector $\gamma_t$ such that
\[ x_t = V_s F_{1,s,n} \gamma_t | x_t + V_s G_{1,s,n} \eta_t + V_s \eta_t. \]  
(A.5)

Furthermore, there exists an $\ell$ dimensional vector $\lambda_t$ with $E_{L-1} \lambda_t = 0$ such that
\[ \eta_t = -(F_{2,s,n} \gamma_t | x_t + V_s G_{2,s,n} \eta_t + \lambda_t). \]  
(A.6)

Eqs. (A5) and (A6) can be thought of as defining $\gamma_t$ and $\lambda_t$. Using Eqs. (A2) and (A4) and the fact that $V_{n-1} V^*_n \gamma_t = x_t - x_t$, we have
\[ 0 = A(s_t) x_t - (B(s_t) x_{t-1} + \Psi(s_t) \eta_t + \Pi \eta_t = [A(s_t) V_s, \Pi] \begin{bmatrix} \gamma_t \\ \lambda_t \end{bmatrix}. \]  
(A.7)

Thus we have shown that any solution of Eq. (1) is of the form defined by Eqs. (A2) through (A7). Though we do not use it, it is also easy to show that any process $(x_t, \eta_t)_{1}^{\infty}$ defined by Eqs. (A2) through (A7) will be a solution of Eq. (1) whose span, conditional on $s_t = i$, is the column space of $V_s$.

If the process $(x_t, \eta_t)_{1}^{\infty}$ is a MSV solution, then it must be the case that $k_t \geq n - \ell$ because there can be no restrictions on $y_t$ and $z_t$. Also, the matrix $[A(s_t) V_s, \Pi]$ must be of full column rank because $\gamma_t$, and hence $\lambda_t$, must be zero. So, the matrix $[A(s_t) V_s, \Pi]$ is invertible. This implies that we can define
\[ \begin{bmatrix} F_{1,i,j} \\ F_{2,i,j} \end{bmatrix} = [A(i) V, \Pi]^{-1} B(i) \]
and
\[ \begin{bmatrix} F_{1,i,j} \\ F_{2,i,j} \end{bmatrix} = \begin{bmatrix} F_{1,i,j} \\ F_{2,i,j} \end{bmatrix} \end{bmatrix} V_j. \]

Substituting this into Eq. (A.3) gives Eq. (9) and substituting this into Eqs. (A5) and (A6) and using the fact that $V_{n-1} V^*_n x_{n-1} = x_{n-1}$ gives Eqs. (5) and (6). This completes the proof of Theorem 1. □

Appendix B. Singular $A(i)$

Using the notation of Section 2, we know that
\[ A(i) V_i = \begin{bmatrix} I_{n-\ell} \\ -X_i \end{bmatrix}. \]  
(A.8)

If $A(i)$ were non-singular, then Eq. (A.8) is easily solved and the results of Section 2 follow. We now consider the case in which $A(i)$ may be singular. We can use the QR decomposition to find an invertible matrix $U_i$ such that $A(i) U_i$ is of the form
\[ \begin{bmatrix} I_{n-\ell} \\ 0_{n-\ell,\ell} \\ C_{l,1} \\ C_{l,2} \end{bmatrix}. \]

Note that $C_{l,1}$ is $\ell \times n - \ell$ and $C_{l,2}$ is $\ell \times \ell$. If the QR decomposition of $A(i)$ is
\[ A(i) = QR_i = Q_i \begin{bmatrix} R_{1,1} \\ R_{2,1} \\ 0_{\ell,n-\ell} \\ R_{2,3} \end{bmatrix}, \]
where $R_{1,1}$ is $n - \ell \times n - \ell$, $R_{2,1}$ is $n - \ell \times \ell$, and $R_{2,3}$ is $\ell \times \ell$, then
\[ U_i = Q_i \begin{bmatrix} (R_{1,1})^{-1} \\ 0_{\ell,n-\ell} \\ I_{\ell} \end{bmatrix}. \]
is the required matrix. The matrix $R_{i,1}$ is invertible because the matrix $a_i(t)$ has full row rank. Eq. (A.8) implies that

$$U_i^{-1}V_i = \begin{bmatrix} I_{n-\ell} \\ -Z_j \end{bmatrix}$$

for some $\ell \times n - \ell$ matrix $Z_j$ and that $X_i = C_{i,2}Z_i - C_{i,1}$. Substituting this into Eq. (9), we obtain

$$\sum_{i=1}^h p_{ij}[C_{i,2}Z_i - C_{i,1}] \Sigma_j B(i)U_i \begin{bmatrix} I_{n-\ell} \\ -Z_j \end{bmatrix} = O_{r,n-\ell}.$$ 

Let $Z = (Z_1, \ldots, Z_h)$, define $g_j$ to be the function from $R_t^{h(n-\ell)}$ to $R_t^{(n-\ell)}$ given by

$$g_j(Z) = \sum_{i=1}^h p_{ij}[C_{i,2}Z_i - C_{i,1}] \Sigma_j B(i)U_i \begin{bmatrix} I_{n-\ell} \\ -Z_j \end{bmatrix} = 0_{r,n-\ell},$$

and $g$ to be the function from $R_t^{h(n-\ell)}$ to $R_t^{h(n-\ell)}$ given by

$$g(Z) = (g_1(Z), \ldots, g_h(Z)).$$

We now have the following algorithm for finding MSV solutions.

**Algorithm 3.** Let $Z^{(0)} = (Z_1^{(1)}, \ldots, Z_h^{(1)})$ be an initial guess. If the $k$th iteration is $Z^{(k)} = (Z_1^{(k)}, \ldots, Z_h^{(k)})$, then the $(k+1)$th iteration is given by

$$vec(Z^{(k+1)}) = vec(Z^{(k)}) - g'(Z^{(k)})^{-1}vec(g(Z^{(k)})).$$

where

$$g'(Z) = \begin{bmatrix} \frac{\partial g_1(Z)}{\partial Z_1} & \cdots & \frac{\partial g_1(Z)}{\partial Z_h} \\
\vdots & \ddots & \vdots \\
\frac{\partial g_h(Z)}{\partial Z_1} & \cdots & \frac{\partial g_h(Z)}{\partial Z_h} \end{bmatrix}.$$ 

The sequence $Z^{(k)}$ converges to a root of $g(Z)$.

As before, it is straightforward to verify that for $i \neq j$,

$$\frac{\partial g_j}{\partial Z_i}(Z) = p_{ij} \left( I_{n-\ell} - 0_{n-\ell}, Z_j \right)' \otimes C_{i,1}$$

and for $i = j$,

$$\frac{\partial g_j}{\partial Z_j}(Z) = p_{ij} \left( I_{n-\ell} - 0_{n-\ell}, Z_j \right)' \otimes C_{i,1} + I_{n-\ell} \otimes \left( \sum_{k=1}^h p_{kj}[C_{k,1}Z_k + C_{k,2}] \Sigma_j B(k)U_k \begin{bmatrix} I_{n-\ell} \\ -I_j \end{bmatrix} \right).$$

References


