1 Computation of the Likelihood Function

1.1 Computing the Observable Structure

Each of the stochastic linear rational expectations models that we consider can be cast in the format

$$\sum_{i=-\tau}^{0} H_i x_{t+i} + \sum_{i=1}^{\theta} H_i E_t(x_{t+i}) = \epsilon_t \,, \tag{1}$$

where, as in section whatever τ and θ are positive integers, x_t is a vector of variables, and the H_i are conformable n-square coefficient matrices, where n is the number of endogenous variables in the model. The coefficient matrices H_i are completely determined by a set of underlying structural parameters Θ. The expectation operator $E_t(\cdot)$ denotes mathematical expectation conditioned on the process history through period t,[1]

$$E_t(x_{t+i}) = E(x_{t+i} | \, x_t, x_{t-1}, \ldots) \,.$$

The random shock ϵ_t is independently and identically distributed $N(0, \Omega)$. Note that the covariance matrix Ω is singular whenever equation (1) includes identities. Of little importance are accounting identities such as the national income accounting identity linking GDP and consumption, investment, government expenditures, and net exports. Of more importance are "expectational identities" such as the identity that defines the *ex ante* long-term real interest rate, ρ, in the pure expectations hypothesis definition of the long rate

$$\rho_t = \sum_{i=0}^{\infty} \beta^i E_t(r_{t+i} - \pi_{t+i+1})$$

[1]The code for computing the observable structure allows an expectations viewpoint date of either t or $t-1$. For simplicity, we focus on the t-period expectations case here.

Expectational identities are important because they define variables such as the long real rate that can only be observed within the context of the model.

Because ϵ_t is white noise, $E_t(\epsilon_{t+k}) = 0$ for $k > 0$. Leading equation (1) by one or more periods and taking expectations conditioned on period-t information yields a deterministic forward-looking equation in expectations,

$$\sum_{i=-\tau}^{\theta} H_i E_t(x_{t+k+i}) = 0, \qquad k > 0\,. \tag{2}$$

We use the AIM procedure detailed above to solve equation (2) for expectations of the future in terms of expectations of the present and the past. For a given set of initial conditions, $\{E_t(x_{t+k+i}) \,:\, k > 0,\, i = -\tau, \ldots, -1\}$, if equation (2) has a unique solution that grows no faster than a given upper bound, that procedure computes the vector autoregressive representation of the solution path,

$$E_t(x_{t+k}) = \sum_{i=-\tau}^{-1} B_i E_t(x_{t+k+i}), \qquad k > 0\,. \tag{3}$$

In the models we consider here, the roots of equation (3) lie on or inside the unit circle.

Using the fact that $E_t(x_{t-k}) = x_{t-k}$ for $k \geq 0$, equation (3) is used to derive expectations of the future in terms of the realization of the present and the past. These expectations are then substituted into equation (1) to derive a representation of the model that we call the *observable structure*,

$$\sum_{i=-\tau}^{0} S_i x_{t+i} = \epsilon_t\,. \tag{4}$$

Equation (4) is a structural representation of the model because it is driven by the structural disturbance, ϵ_t; the coefficient matrix S_0 contains the contemporaneous relationships among the elements of x_t. It is an observable representation of the model because it does not contain unobservable expectations.

1.2 Computing the Likelihood Function

Having obtained the observable structure of equation 4, it is relatively straightforward to compute the value of the likelihood function given the data and parameter values. The likelihood is defined as

$$\mathcal{L} = T(\log|\mathcal{J}| - .5\log|\hat{\Omega}|) \tag{5}$$

where T is the sample size, $\mathcal{J}$ is the Jacobian of transformation (which is time-invariant by assumption), and Ω is the variance-covariance matrix of the structural residuals ϵ_t.

Consider concatenating the n by n coefficient matrices S_i, ordered left to right from $i = -\tau$ to 0. We denote this n by $n \times (\tau + 1)$ matrix `SCOF`. Define the vector stack of the endogenous variables at time t as $X_t = [x_{t-\tau}, \ldots, x_t]'$. Thus equation 4 may be rewritten

$$\texttt{SCOF} X_t = \epsilon_t$$

In computing the value of the likelihood, it will be useful to partition `SCOF` as follows. Denote stochastic equations by the subscript s, identity equations by the subscript i, and denote data variables with the subscript d, and "not-data" variables (such as the unobserved long real rate defined above) with the subscript n. Arbitrarily ordering the observable structure so that stochastic equations appear in the top rows and data variables in the left columns of each block, we can write equation 4 as

$$\left[\begin{array}{c|cc} S_{s,t-1} & S_{s,d} & S_{s,n} \\ \hline S_{i,t-1} & S_{i,d} & S_{i,n} \end{array}\right] \left[\begin{array}{c} X_{t-1} \\ \\ \hline X_{d,t} \\ X_{n,t} \end{array}\right] = \left[\begin{array}{c} \epsilon_t \\ \hline 0 \end{array}\right] \tag{6}$$

$S_{s,t-1}$ denotes the coefficient block of `SCOF` for the lagged variables that enter the stochastic equations; $S_{i,t-1}$ is the corresponding block for identity equations. The right-hand-most n by n block of equations, representing the coefficients on contemporaneous variables, is further partitioned vertically into its data and not-data components.

For each observation t, we use this concatenated, partitioned version of the observable structure to solve for the residuals ϵ_t. First, solve for the period-t not-data variables as

$$X_{n,t} = -S_{i,n}^{-1}[S_{i,t-1}X_{t-1} + S_{i,d}X_{d,t}] \tag{7}$$

Now substitute the solution for $X_{n,t}$ into the top rows of equation 6 to solve for ϵ_t:

$$\epsilon_t = S_{s,t-1}X_{t-1} + S_{s,d}X_{d,t} - S_{s,n}S_{i,n}^{-1}[S_{i,t-1}X_{t-1} + S_{i,d}X_{d,t}] \tag{8}$$

The residuals for each time period $t = 1, \ldots, T$ are computed, and the residual covariance matrix is then computed as

$$\Omega = (1/T)\epsilon\epsilon' \tag{9}$$

Note that implict in the solution for the residuals (equation 8) is the definition of the Jacobian, $\partial\epsilon_t/\partial X_t$,

$$\mathcal{J} \equiv S_{s,d} - S_{s,n}S_{i,n}^{-1}S_{i,d} \tag{10}$$

2 Full Information Maximum Likelihood Estimation

Maximum likelihood estimation consists of finding the parameter values Θ, implicit in the coefficient matrices H_i of equation 1, that maximize equation 5. We use Matlab's sequential quadratic programming algorithm `constr` to maximize the likelihood function, subject to several types of constraints:

1. Parameter boundary constraints (upper and lower bounds for the elements of Θ);

2. Equality constraints of the form $F(\Theta) = 0$;

3. Inequality constraints of the form $G(\Theta) \leq 0$. Our routine always enforces the nonlinear inequality constraint that the current parameter setting must be consistent with the correct number of large roots (the number of roots whose magnitude exceeds the specified upper bound is consistent with a unique, stable solution) in a converged solution.

The procedure uses numerical derivatives of the likelihood with respect to the parameters and of the constraints with respect to the parameters. Standard errors are computed from the numerical estimate of the Hessian, $H = \partial^2 \mathcal{L} / \partial \Theta^2$ as

$$se = \sqrt{(}diag(H^{-1})) \tag{11}$$

where $diag$ indicates the diagonal elements of the inverse Hessian matrix.