ESTIMATION AND SOLUTION OF LINEAR RATIONAL EXPECTATIONS MODELS USING A POLYNOMIAL MATRIX FACTORIZATION

C. Ates DAGLI

SPSS Inc., Chicago, IL 60611, USA

John B. TAYLOR

Stanford University, Stanford, CA 94305, USA

A method to solve and estimate multivariate linear rational expectations models is described. The method is based on an iterative factorization of the polynomial matrix that describes the lags and expected leads in the model. Our experience is that the method works well in a variety of applications where other methods are either difficult or expensive to use.

This note describes a method that can be used to solve and estimate multivariate linear rational expectations models. A number of methods to solve and estimate linear rational expectations models have been described in the literature [e.g., Blanchard and Kahn (1980), Chow (1983), Fair and Taylor (1983), Hansen and Sargent (1981), Taylor (1980), and Wallis (1980)]. The method described here is a generalization of the factorization techniques used in Hansen and Sargent (1981) and Taylor (1980) to general non-symmetric polynomial matrices. We have had considerable success during the last few years using the method on linear models where some of the other methods appear to us to be difficult to apply or expensive to use. Although we do not have detailed information about the relative advantages and disadvantages of this method compared to the other methods, our experience indicates that a description of the method may be of interest to others studying the quantitative implications of rational expectations. The computer routines necessary to use the method have been programmed as part of a general multivariate time series computer package and are fairly easy to use.

We consider the following linear rational expectations model:

$$B_0 y_t + B_1 y_{t-1} + \cdots + B_p y_{t-p} + A_0 \hat{y}_t + \cdots + A_q \hat{y}_{t+q} = \Gamma x_t + u_t, \quad (1)$$

where the matrices A_i and B_i are $n \times n$, y_t is an n-dimensional vector of endogenous variables at time t, x_t is a k-dimensional vector of exogenous variables at time t, Γ is an $n \times k$ matrix of parameters, and u_t is an n-dimensional vector of disturbances which is not generally serially independent. The hat over a variable indicates its conditional expectation based on information through time $t - 1$. (The method can also be used for the case where the conditioning set includes information through time t.)

The error term u_t is assumed to follow a stationary stochastic process which we represent in the general linear form as

$$u_t = \Delta(L)\varepsilon_{1t}, \quad (2)$$

where ε_{1t} is an independent and identically distributed random vector with mean 0, and $\Delta(L)$ is a matrix polynomial in the lag operator L. In order to compute future expectations it is necessary to postulate a stationary stochastic process for the exogenous variables which we represent as

$$x_t = \Lambda(L)\varepsilon_{2t}, \quad (3)$$

where ε_{2t} is an independent and identically distributed random vector with mean 0, and $\Lambda(L)$ is a matrix polynomial in the lag operator L. Agents described by the model are assumed to know $\Lambda(L)$ and past realizations of x_t and y_t at time t; they are assumed not to know the future realizations of x_t at time t. We assume for the purposes of forming the likelihood function for estimation that the random vector $\varepsilon_t' = (\varepsilon_{1t}', \varepsilon_{2t}')$ is normally distributed with mean zero and covariance matrix Ω. We also use the normalization that Λ_0 and Δ_0 are equal to identity matrices.

The solution method gives the coefficients (in numerical form) of a constrained simultaneous vector autoregressive moving average (VARMA) model which is equivalent to (1) but in which expectations do not appear. This VARMA model can then be used for simulations or full information maximum likelihood estimation. The VARMA model is assumed to be stationary and is given by

$$B_0 y_t = C(L) y_t + D(L) x_t + R(L) \varepsilon_{1t}, \quad (4)$$

where

$$C(L) = B_0 \theta_0^{-1} (\theta_0 - \Theta(L)), \quad (5)$$

$$D(L) = \left(B_0 \theta_0^{-1} \psi_2(L) + \Gamma \right) \Lambda^{-1}(L), \quad (6)$$

$$R(L) = \left(B_0 \theta_0^{-1} \psi_1(L) + I \right), \quad (7)$$

where $C(L)$, $D(L)$, and $R(L)$ are matrix polynomials in the lag operator L which are defined in terms of three other matrix polynomials in the lag operator: $\psi_1(L)$, $\psi_2(L)$ and $\theta(L)$ with zero-order coefficient matrix θ_0. These latter three polynomials are defined as follows: Let

$$H(L) = B_1 L + \cdots + B_p L^p + (B_0 + A_0) + A_1 L^{-1} + \cdots + A_q L^{-q}. \quad (8)$$

Assume that there is a unique factorization of $H(L)$,

$$H(L) = \phi(L^{-1})\theta(L), \quad\quad\quad\quad (9)$$

where

$$\Phi(L^{-1}) = I + \cdots + \Phi_q L^{-q} \quad \text{and} \quad \theta_0(L) = \theta_0 + \cdots + \theta_p L^p$$

are both real where the roots of the determinental polynomials of $\theta(L)$ and $\Phi(L)$ have all roots outside the unit circle. Then the matrix polynomial $\theta(L)$ in (5), (6) and (7) is given by (9) and $\Psi_1(L)$ and $\Psi_2(L)$ are given by

$$\Psi_1(L) = \left[\left(\Phi(L^{-1}) \right)^{-1} \Delta(L) \right]^+, \quad\quad\quad\quad (10)$$

$$\Psi_2(L) = \left[\left(\Phi(L^{-1}) \right)^{-1} \Gamma\Lambda(L) \right]^+, \quad\quad\quad\quad (11)$$

where $[\cdot]^+$ denotes that all terms involving non-positive powers of L have been dropped from the expression inside the brackets.

That model (4) is a characterization of (1) is discussed in Whiteman (1983). Model (4) is a unique stationary characterization of (1) if the factorization (9) is unique. In what follows we assume that (4) is unique, and focus on the problem of obtaining (4) numerically.

The main computational problem in going from (1) to (4) is the matrix polynomial factorization in (9). Evaluation of (10) and (11) is straightforward in the case where $\Delta(L)$ and $\Lambda(L)$ are finite order, and the methods of Hansen and Sargent (1981) can be used to evaluate (10) and (11) when $\Delta(L)$ and $\Lambda(L)$ are infinite order.

The algorithm for factoring (9) that is described here is an interative one and does not require calculation of the characteristic roots of the determinental polynomials. In special cases more direct non-iterative methods might be used. Our experience is that the iterative procedure converges quickly, though no formal proof of convergence is available.

Eq. (9) implies a set of identities in the coefficients which can be obtained by equating coefficients of like-powered lag operators on both sides of the equation. These identities are

$$H_{-k} = \sum_{i=0}^{\min(p,q-k)} \Phi_{k+i}\theta_i, \qquad k = 0,\ldots,q, \tag{12}$$

$$H_k = \sum_{i=0}^{\min(q,p-k)} \phi_i\theta_{k+i}, \qquad k = 0,\ldots,p, \tag{13}$$

where we have included one of the identitites ($k = 0$) twice for ease of exposition below.

Note that, if $\theta(L)$ is given, the Φ_k can be computed using (12) as follows:

$$\phi_k = \left[H_{-k} - \sum_{i=1}^{\min(p,q-k)} \Phi_{k+i}\theta_i \right]\theta_0^{-1}, \qquad k = q, q-1,\ldots,0.$$

Similarly, if $\Phi(L)$ is given, the θ_k can be computed recursively using (13). The algorithm exploits this form of the equations to perform a series of successive approximations, starting with a guess of one polynomial, computing the second, using the second to get a new guess for the first, and so on. Let $\theta^{(j)}(L)$ be the value of $\theta(L)$ on iteration j and let $\Phi^{(j)}(L)$ be the value of $\Phi(L)$ on iteration j. The algorithm can then be described in the following steps:

(a) Pick a starting value $\theta^{(0)}(L)$ for $\theta(L)$ on the first iteration.

(b) Computer the $\Phi_k^{(j)}$ using (12) with $\theta(L) = \theta^{(j-1)}(L)$ starting with $j = 1$.

(c) Computer the $\theta_k^{(j)}$ using (13) with $\Phi(L) = \Phi^{(j)}(L)$.

(d) If the coefficients $\theta_k^{(j)}$ and $\theta_k^{(j-1)}$, $k = 0,\ldots,p$, are within a specified tolerance range of each other, then stop. Otherwise repeat starting from step b, increasing j by 1.

Note that if $\Phi(L^{-1})$ and $\theta(L)$ constitute a factorization of $H(L)$, so do $\hat{\Phi}(L^{-1}) = \Phi(L^{-1})C$ and $\hat{\theta}(L) = C^{-1}\theta(L)$ for any non-singular matrix C. We adopt the normalization convention that $\Phi_0 = I$ and impose it on the result of the factorization. Thus, if the factorization converges at the mth iteration, we use

$$\Phi(L) = \Phi^{(m)}(L)\left[\Phi_0^{(m)}\right]^{-1} \quad \text{and} \quad \theta(L) = \left[\Phi_0^{(m)}\right]\theta^{(m)}(L).$$

We have no theoretically-based guides for choosing a starting value. However, we have had considerable experimental success with the simple choice $\theta^{(0)}(L) = H_0 + H_1L + \cdots + H_pL^P$, or equivalently, with $\Phi^{(0)}(L) = I$.

When this factorization method is used during the numerical maximization of the likelihood function, the algorithm is called upon to factor a sequence of $H(L)$'s each differing slightly from the previous one. In such cases, the choice of the $\theta(L)$ from the previous factorization as the starting value in the current one has proved to be extremely effective in reducing the number of iterations.

It is easily seen that when a row of $H(L)$ has no forward terms, neither does the corresponding row of $\Phi(L^{-1})$. This result, coupled with the observation that most rational expectations models have forward components in only a subset of the equations, allows us to economize further on computation.

An alternative characterization of the algorithm may be given in terms of the formal long division of matrix polynomials. For any matrix polynomial $P(L)$ of degree n, define

$$\tilde{P}(L) = L^n P(L^{-1}) = P_0 L^n + \cdots + P_n,$$

and note that the non-zero roots of $|\tilde{P}(L)|$ are the reciprocals of the non-zero roots of $|P(L)|$. Let

$$E(L) = L^q H(L) = E_0 + E_1 L + \cdots + E_{p+q} L^{p+q}.$$

Thus

$$E(L) = L^q \Phi(L^{-1})\theta(L) = \tilde{\Phi}(L)\theta(L).$$

The factorization problem may now be recast as the problem of factoring $E(L)$ into $\tilde{\Phi}(L)$ of degree q with all its roots inside the unit circle and $\theta(L)$ of degree p with all its roots outside the unit circle. Step (b) of the algorithm is equivalent to the right-division of $\hat{E}(L)$ by $\tilde{\theta}^{(j-1)}(L)$ to obtain $\Phi^{(j)}(L)$ as the quotient; that is $\tilde{E}(L) = \Phi^{(j)}(L)\tilde{\theta}^{(j-1)}(L) + \text{remainder}$. Analogously, step (c) is the left-division of $E(L)$ by $\tilde{\Phi}^{(j)}(L)$ to give $\theta^{(j)}(L)$ as the quotient; that is, $E(L) = \tilde{\Phi}^{(j)}(L)\theta^{(j)}(L) + \text{remainder}$.

As a check on the correctness of the factorization, we recommend that the locations of the zeroes of $|\Phi(L)|$ and $|\theta(L)|$ relative to the unit circle be determined. Methods to perform this task without having to compute the zeroes themselves exist [see Marden (1966)], and have been implemented as computer programs by the authors.

An example of how eqs. (12) and (13) are solved in each iteration for particular values of p and q may be helpful. Suppose that $p = 3$ and $q = 2$. In

this case, eqs. (12) and (13) become

$$H_{-2} = \Phi_2 \theta_0,$$

$$H_{-1} = \Phi_1 \theta_0 + \Phi_2 \theta_1,$$ (12')

$$H_0 = \Phi_0 \theta_0 + \Phi_1 \theta_1 + \Phi_2 \theta_2,$$

$$H_3 = \Phi_0 \theta_3,$$

$$H_2 = \Phi_0 \theta_2 + \Phi_1 \theta_3,$$

$$H_1 = \Phi_0 \theta_1 + \Phi_1 \theta_2 + \Phi_2 \theta_3,$$ (13')

$$H_0 = \Phi_0 \theta_0 + \Phi_1 \theta_1 + \Phi_2 \theta_2.$$

The objective is to solve this set of equations for the matrices θ_0, θ_1, θ_2, θ_3, Φ_1 and Φ_2. Given a guess of Φ_0, Φ_1 and Φ_2 (such as $\Phi_0 = I$ and $\Phi_1 = \Phi_2 = 0$), it is clear how one can compute θ_0, θ_1, θ_2, and θ_3, using eq. (13'). Moreover, given computed values for these θ matrices, a new set of guesses for Φ_0, Φ_1 and Φ_2 can be calculated using eq. (12'). These new values of Φ_0, Φ_1 and Φ_2 can in turn be used to start a new iteration again by computing a new set of θ coefficients, and so on.

The above approach to estimating rational expectations models has been applied in several empirical studies with apparent success. A macroeconometric model of the U.S. similar to that described in Taylor (1979) was estimated using full information maximum likelihood and the factorization procedure. The model contains five endogenous variables, lagged dependent variables in four of the five equations, and builds up from a full IS-LM structure so that the expected inflation rate appears (via the expectation of next period's price level) in the aggregate demand equation. The model also contains anticipatory wage behavior so that expected future wages and output appear in the wage setting equation. The resulting model was therefore of the form of (1) with $p = 7$ (seven leads of wages and unemployment in the wage equation, and one lead of price in the demand equation) and $q = 7$. The model was then estimated using the procedure outlined above. An open economy macroeconometric model of the U.S. and Germany was estimated by Rehm (1982) using full information maximum likelihood. The model contains eight equations with five leads and five lags ($p = q = 5$). The factorization method was also used for solving the model and characterizing its dynamic properties.

Other applications of the factorization method include work by Johnson (1982) and Abraham (1983). Johnson's application is to a two-country rational expectations model, which can be reduced to a two-dimensional, two-sided

matrix polynomial with $p = q = 1$. The factorization algorithm was used to solve (but not to estimate) the model. In the application of Abraham (1983), model (1) is derived from a formal utility maximization problem along the lines of Hansen and Sargent (1981). However, because the matrix equations do not satisfy certain non-singularity conditions required for the application of the Hansen–Sargent method, he uses the factorization algorithm reported in this paper. The method seems to work successfully in that the roots of the factored polynomials are on the correct side of the unit circle.

In addition to these applications, we can report on experiments with 40 test polynomials to which we applied the algorithm. These test polynomials were constructed by multiplying $\Phi(L^{-1})$ and $\theta(L)$ pairs with the appropriate root properties to form the $H(L)$'s. Thus, in each case, we knew the correct solution to the factorization problem. The $\Phi(L^{-1})$ and $\theta(L)$, in turn, were built up as the product of factors of the form $I + D_i L$, with the D_i generated pseudo-randomly and scaled so as to have all their characteristic roots inside the unit circle. This ensures that the zeroes of $I + D_i L$ all lie outside the unit circle. The forty polynomials ranged from 2×2 systems with $p = q = 1$ to 9×9 systems with $p = q = 6$. They were not designed to correspond to any economic application. The algorithm converged in all cases. In one case (a 3×3 system with $p = 4$ and $q = 9$) the converged values of $\Phi(L^{-1})$ and $\theta(L)$ did not equal the correct solution. In this case $\Phi(L^{-1})\theta(L)$ was quite different from the true $H(L)$; in practice such an incorrect solution could be detected by a procedure which checked that the converged values satisfy the original polynomial equations $\Phi(L^{-1})\theta(L) = H(L)$ to within a given degree of accuracy.

The timings for convergence of the forty test cases (on an IBM 4341 Model Group 2 computer) varied from about 0.01 second for the 2×2 system to 1.9 seconds for the 9×9 system.

References

Abraham, J., 1983, Wage determination under long-term labor contracts, Ph.D. dissertation (Princeton University, Princeton, NJ).

Blanchard, O. and C. Kahn, 1980, The solution of linear difference models under rational expectations, Econometrica 48, 1305–1311.

Chow, G.C., 1983, Econometrics (McGraw-Hill, New York).

Fair, R.C. and J.B. Taylor, 1983, Solution and estimation of dynamic non-linear rational expectations models, Econometrica 51, 1169–1185.

Hansen, L.P. and T.J. Sargent, 1981, Linear rational expectations models for dynamically interrelated variables, in: R.E. Lucas, Jr. and T.J. Sargent, eds., Rational expectations and econometric practice (University of Minnesota Press, Minneapolis, MN).

Johnson, R., 1982, Monetary stabilization and interdependence in a two country model, Unpublished paper (Princeton University, Princeton, NJ).

Marden, M., 1966, Geometry of polynomials (American Mathematical Society, Providence, RI).

Rehm, D.E., 1982, Staggered contracts, capital flows, and macroeconomic stability in the open economy, Ph.D. dissertation (Columbia University, New York).

Taylor, J.B., 1977, Conditions for unique solutions in stochastic macroeconomic models with rational expectations, Econometrica 45, 1377–1385.

Taylor, J.B., 1979, An econometric business cycle model with rational expectations: Some estimation results, Unpublished paper.

Taylor, J.B., 1980, Output and price stability: An international comparison, Journal of Economic Dynamics and Control 2, 109–132.

Wallis, K.F., 1980, Econometric implications of the rational expectations hypothesis, Econometrica 48, 48–74.

Whiteman, C.H., (1983), Linear rational expectations models: A user's guide (University of Minnesota Press, Minneapolis, MN).