TOPICS IN MACROECONOMICS: MODELLING INFORMATION, LEARNING AND EXPECTATIONS

SOLVING LINEAR RATIONAL EXPECTATIONS MODELS

KRISTOFFER P. NIMARK

Three ways to solve a linear model

Solving a model using full information rational expectations as the equilibrium concept involves integrating out expectations terms from the structural equations of the model by replacing agents' expectations with the mathematical expectation, conditional on the state of the model.[1] These notes describes three different ways of doing this. The first method, which is the standard method for solving more elaborate (linear) models, is to decouple the stable and unstable dynamics of the model and set the unstable part to zero. The second method, the method of undetermined coefficients, can be very quick when feasible and illustrates the fixed point nature of the rational expectations solution. The third method is to integrate out expectations by replacing them with linear projections on observable variables. This is the method that has been used to solve some imperfect information models, e.g. Townsend (1983), Singleton (1987), Sargent (1991) and Allen, Morris and Shin (2006). In order to understand its uses and limitations for solving imperfect information models, it is helpful to first understand how the method of projections on observable variables works in a full information model. The Appendix briefly introduces the Projection Theorem.

As a vehicle to demonstrate the different solution methods, we will use a simple New-Keynesian model

Date: November 2, 2010.

[1]See Muth (1961).

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(y_t - \overline{y}_t) \tag{0.1}$$

$$y_t = E_t y_{t+1} - \sigma\left(i_t - E_t \pi_{t+1}\right) \tag{0.2}$$

$$i_t = \phi \pi_t \tag{0.3}$$

$$\overline{y}_t = \rho \overline{y}_{t-1} + u_t \tag{0.4}$$

where π_t, y_t, y_t, i_t are inflation, output, potential output and nominal interest rate respectively. This model has a single variable, potential output $\overline{y}_t$, as the state.

1. Stable/unstable decoupling

This method is originally due to Blanchard and Kahn (1980) but the computational aspects of the method has been further developed by others, for instance Klein (2000). The most accessible reference is probably Soderlind (1999), who also has code posted on his web site. The method has several advantages: Not only does it deliver a solution relatively fast, it also provides conditions for when a solution exists and when the solution is unique.

Start by putting the model (0.1) -(0.4) into matrix form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \beta & 0 \\ 0 & \sigma & 1 \end{bmatrix} \begin{bmatrix} \overline{y}_{t+1} \\ E_t \pi_{t+1} \\ E_t y_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ \kappa & 1 & -\kappa \\ 0 & \sigma\phi & 1 \end{bmatrix} \begin{bmatrix} \overline{y}_{t-1} \\ \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u_{t+1} \tag{1.1}$$

or

$$A_0 \begin{bmatrix} x^1_{t+1} \\ E_t x^2_{t+1} \end{bmatrix} = A_1 \begin{bmatrix} x^1_t \\ x^2_t \end{bmatrix} + C_1 u_{t+1} \tag{1.2}$$

where x^1_t is vector containing the pre-determined and/or exogenous variables (i.e. $\overline{y}_t$) and x^2_t a vector containing the forward looking (”jump”) variables (i.e. $E_t y_{t+1}$ and $E_t \pi_{t+1}$). Pre-multiply both sides by A_0^{-1} to get

$$\begin{bmatrix} x^1_{t+1} \\ E_t x^2_{t+1} \end{bmatrix} = A \begin{bmatrix} x^1_t \\ x^2_t \end{bmatrix} + C u_{t+1} \tag{1.3}$$

where $A = A_0^{-1}A_1$ and $C = A_0^{-1}C_1$. For the model to have unique stable solution the number of stable eigenvalues of A must be equal to the number of exogenous/pre-determined variables. Use a Schur decomposition to get

$$A = ZTZ^H \tag{1.4}$$

where T is (at least) upper block triangular

$$T = \begin{bmatrix} T_{11} & T_{12} \\ \mathbf{0} & T_{22} \end{bmatrix} \tag{1.5}$$

and Z is a unitary matrix so that $Z^H Z = ZZ^H = I$ ($\Longrightarrow$ $Z^H = Z^{-1}$). (For any square matrix W, $W^{-1}AW$ is a so called similarity transformation of A. Similarity transformations has the property that they do not change the eigenvalues of a matrix, so $T(= Z^H AZ)$ has the same eigenvalues as A and this would be true even if Z was not unitary.) We can always choose Z and T so that the unstable eigenvalues of A are shared with T_{22}, which turns out to be useful.

Define the auxiliary variables

$$\begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} = Z^H \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} \tag{1.6}$$

We can then rewrite the system (1.3) as

$$Z^H \begin{bmatrix} x_{t+1}^1 \\ E_t x_{t+1}^2 \end{bmatrix} = Z^H ZTZ^H \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} \tag{1.7}$$

or equivalently

$$E \begin{bmatrix} \theta_{t+1} \\ \delta_{t+1} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ \mathbf{0} & T_{22} \end{bmatrix} \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} \tag{1.8}$$

since $Z^H Z = I$. For this system to be stable, the auxiliary variables associated with the unstable roots in T_{22} must be zero for all t. Imposing $\delta_t = 0 \forall t$ reduces the system to

$$\theta_t = T_{11}\theta_{t-1}$$

To get back the original variables we simply use that

$$\begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} = \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} \theta_t \tag{1.9}$$

or

$$\begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} = \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} Z_{11}^{-1} x_t^1$$

which is the solution to the model. It is in the form

$$x_t^1 = M x_{t-1}^1 + \varepsilon_t \tag{1.10}$$

$$x_t^2 = G x_t^1 \tag{1.11}$$

where $M = Z_{11}T_{11}Z_{11}^{-1}$ (= ρ in our example) and $G = Z_{21}Z_{11}^{-1}$.

2. Method of undetermined coefficients

The method of undetermined coefficients is quick when feasible and illustrates well the fixed point nature of rational expectations equilibria. Since we know that the state of the model (0.1) -(0.4) is the exogenous potential output, we can conjecture a solution of the model in the following form (indeed, it is the same form as the solution of the model above delivered).

$$\overline{y}_t = \rho \overline{y}_{t-1} + u_t \tag{2.1}$$

$$\pi_t = a \overline{y}_t \tag{2.2}$$

$$y_t = b \overline{y}_t \tag{2.3}$$

Both inflation and output are linear functions of the state. Solving the model implies finding the coefficients a and b. Start by substituting in the conjectured solution into the structural equations (0.1) -(0.4) so that

$$a\overline{y}_t = \beta a\rho\overline{y}_t + \kappa\left(b\overline{y}_t - \overline{y}_t\right) \tag{2.4}$$

$$b\overline{y}_t = b\rho\overline{y}_t - \sigma\left[\phi a\overline{y}_t - a\rho\overline{y}_t\right] \tag{2.5}$$

where we also used that $i_t = \phi a\overline{y}_t$. Equating coefficients on the LHS and the RHS we get

$$a - \beta a\rho - \kappa b = -\kappa \tag{2.6}$$

$$b - b\rho + \sigma\phi a - \sigma a\rho = 0 \tag{2.7}$$

which is a system of linear equations in a and b

$$\begin{bmatrix} 1-\beta\rho & -\kappa \\ \sigma\phi - \sigma\rho & 1-\rho \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -\kappa \\ 0 \end{bmatrix} \tag{2.8}$$

which can be solved by pre multiplying both sides with the inverse of the coefficient matrix on the LHS

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1-\beta\rho & -\kappa \\ \sigma\phi - \sigma\rho & 1-\rho \end{bmatrix}^{-1}\begin{bmatrix} -\kappa \\ 0 \end{bmatrix} \tag{2.9}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -\kappa\frac{\rho-1}{\rho+\beta\rho-\beta\rho^2-\kappa\sigma\phi+\kappa\sigma\rho-1} \\ -\kappa\frac{\sigma\phi-\sigma\rho}{\rho+\beta\rho-\beta\rho^2-\kappa\sigma\phi+\kappa\sigma\rho-1} \end{bmatrix} \tag{2.10}$$

The vector $\begin{bmatrix} a & b \end{bmatrix}'$ equals the vector G from the stable/unstable eigenvalue decoupling method of Section1 above.

3. Replacing expectations with linear projections on observables

The third method uses that projections of the future values of variables on observables gives optimal expectations (in the sense of minimum error variance) if the observables span the space of the state. In the model (0.1) -(0.4) we can replace $E_t\pi_{t+1}$ and $E_t y_{t+1}$ with linear projections of these variables on current inflation. (There is nothing special about inflation. Projecting onto current output would also work.). We will use that

$$E\left(\pi_{t+1} \mid \pi_t\right) = \frac{cov(\pi_t, \pi_{t+1})}{var\left(\pi_t\right)}\pi_t \quad (3.1)$$

$$E\left(y_{t+1} \mid \pi_t\right) = \frac{cov(\pi_t, y_{t+1})}{var\left(\pi_t\right)}\pi_t \quad (3.2)$$

if the innovations u_t to $\overline{y}_t$ are Gaussian.

Let

$$c_0\pi_t = E^*\left(\pi_{t+1} \mid \pi_t\right) \quad (3.3)$$

$$d_0\pi_t = E^*\left(y_{t+1} \mid \pi_t\right) \quad (3.4)$$

denote initial candidate projections of expected inflation and output on current inflation. We can then write the structural equations (0.1) and (0.2) as

$$\pi_t = \beta c_0\pi_t + \kappa(y_t - \overline{y}_t) \quad (3.5)$$

$$y_t = d_0\pi_t - \sigma\left(\phi\pi_t - c_0\pi_t\right) \quad (3.6)$$

Put the whole system in matrix form

$$\begin{bmatrix} \overline{y}_t \\ \pi_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \kappa & 1-\beta c_0 & -\kappa \\ 0 & -d_0+\sigma\phi-\sigma c_0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \overline{y}_{t-1} \\ \pi_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ \kappa & 1-\beta c_o & -\kappa \\ 0 & -d_0+\sigma\phi-\sigma c_0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u_t \tag{3.7}$$

or

$$X_t = AX_{t-1} + Cu_t$$

The model can be solved by iterating on the following algorithm:

(1) Make an initial guess of c_0 and d_0 in (3.7)

(2) Compute the implied covariances of current inflation and future inflation and output using

$$\begin{aligned} E\left[X_t X_t'\right] &= \Sigma_{XX} \\ \Sigma_{XX} &= A\Sigma_{XX}A' + CC' \end{aligned}$$

and

$$E\left[X_t X_{t+1}'\right] = A\Sigma_{XX}$$

(3) Replace the c_s and d_s with the c_{s+1} and d_{s+1} in ((3.7))

$$\begin{aligned} c_{s+1} &= \frac{cov(\pi_t, \pi_{t+1})}{var\left(\pi_t\right)} \\ d_{s+1} &= \frac{cov(\pi_t, y_{t+1})}{var\left(\pi_t\right)} \end{aligned}$$

using the covariances from Step 2

(4) Repeat Step 2-3 until c_s and d_s converges.

Seems pretty stupid, but it works!

3.1. **When do Solution 3 coincide with Solution 1 and 2?** The process above would need to be amended if the state was of higher dimension. For instance, if we add a "cost push" shock to the system so that

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(y_t - \overline{y}_t) + \varepsilon_t^{\pi}$$

the space of the state would no longer be spanned by the a single variable. We could still use linear projections to solve the model but would need to compute projections as

$$E\left(\begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} \mid \pi_t \, , y_t\right) = D_{\pi y} A \Sigma_{XX} D'_{\pi y} \left(D_{\pi y} \Sigma_{XX} D'_{\pi y}\right)^{-1} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} \tag{3.8}$$

where the $D_{\pi y}$

$$D_{\pi y} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.9}$$

picks out the appropriate covariances. Substituting into the structural equations

$$\pi_t = \beta \mathbf{c}_0 \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} + \kappa(y_t - \overline{y}_t) \tag{3.10}$$

$$y_t = \mathbf{d}_0 \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} - \sigma \left(\phi \pi_t - \mathbf{c}_0 \begin{bmatrix} \pi_t \\ y_t \end{bmatrix}\right) \tag{3.11}$$

where

$$\begin{bmatrix} \mathbf{c}_0 \\ \mathbf{d}_0 \end{bmatrix} = D_{\pi y} A \Sigma_{XX} D'_{\pi y} \left(D_{\pi y} \Sigma_{XX} D'_{\pi y}\right)^{-1} \tag{3.12}$$

This would still give a correct result. However, if we add another shock (or state variable) to the model but continue to assume that potential output is unobservable, the method will no longer produce the same result as the other two methods. The reason is that the space spanned by the observables then do not span the space of the state, so projections on only current inflation and output will not be optimal estimates of the next period values of these variables. In fact, in order to obtain optimal projections given the history of observable variables, it would be necessary to compute the projection of expected inflation and output on the entire history of observable variables. The Kalman filter provides a convenient way of recursively doing exactly that, but without carrying along the complete history of observable variables.

References

[1] Allen, F. S. Morris and H.S., 2006, "Shin, Beauty Contests and Iterated Expectations in Asset Markets", *Review of Financial Studies*, 19, pp719 – 752.

[2] Blanchard, Olivier Jean and Kahn, Charles M, 1980, "The Solution of Linear Difference Models under Rational Expectations", *Econometrica*, vol. 48(5), pages 1305-1.

[3] Brockwell, P.J. and R.A. Davis, 2006, *Time Series: Theory and Methods*, Springer-Verlag.

[4] Klein, P., 2000, "Using the generalized Schur form to solve a multivariate linear rational expectations model", *Journal of Economic Dynamics and Control 24,* pp1405-1423.

[5] Muth, John A. "Rational Expectations and the Theory of Price Movements", *Econometrica* 29, no. 6 (1961): 315-35.

[6] Sargent, Thomas J., 1991, "Equilibrium with Signal Extraction from Endogenous Variables", *Journal of Economic Dynamics and Control* 15, pp245-273.

[7] Singleton, Kenneth J., 1987, "Asset prices in a time series model with disparately informed, competitive traders", in *New Approaches to Monetary Economics,* Eds. W.A. Burnett and K.J. Singleton, Cambridge University Press.

[8] Soderlind, P. 1999, "Solution and estimation of RE models with optimal policy", European Economic Review, Vol 43, pp813-823.

Appendix A. The Projection Theorem

This Appendix defines a few concepts needed for stating the projection theorem. It also relates some of the properties of linear projections on inner product spaces to the solution method described in Section 3 above. For more details and a proof of the projection theorem, see for instance Chapter 2 of Brockwell and Davis (2006).

Definition 1. *(Inner Product Space) An real vector space $\mathcal{H}$ is said to be an inner product space if for each pair of elements x and y in $\mathcal{H}$ there is a number $\langle x, y\rangle$ called the inner product of x and y such that*

$$\langle x, y\rangle = \langle y, x\rangle \tag{A.1}$$

$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \text{ for all } x, y, z \in \mathcal{H} \tag{A.2}$$

$$\langle \alpha x, y\rangle = \alpha \langle x, y\rangle \text{ for all } x, y \in \mathcal{H} \text{ and } \alpha \in \mathbb{R} \tag{A.3}$$

$$\langle x, x\rangle \geq 0 \text{ for all } x \in \mathcal{H} \tag{A.4}$$

$$\langle x, x\rangle = 0 \text{ if and only if } x = \mathbf{0} \tag{A.5}$$

Definition 2. *(Norm) The norm of an element x of an inner product space is defined to be*

$$\|x\| = \sqrt{\langle x, x\rangle} \tag{A.6}$$

Definition 3. *(Cauchy Sequence) A sequence $\{x_n, n = 1, 2, ...\}$ of elements of an inner product space is said to be Cauchy sequence if*

$$\|x_n - x_m\| \to 0 \text{ as } m, n \to \infty$$

i.e. for every $\varepsilon > 0$ there exists a positive integer $N(\varepsilon)$ such that

$$\|x_n - x_m\| < \varepsilon \text{ as } m, n > N(\varepsilon)$$

Definition 4. *(Hilbert Space) A Hilbert space $\mathcal{H}$ is an inner product space which is complete, i.e. every Cauchy sequence $\{x_n\}$ converges in norm to some element $x \in \mathcal{H}$.*

Theorem 1. *(The Projection Theorem) If $\mathcal{M}$ is a closed subspace of the Hilbert Space $\mathcal{H}$ and $x \in \mathcal{H}$, then*
(i) there is a unique element $\widehat{x} \in \mathcal{M}$ such that

$$\|x - \widehat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\|$$

and
(ii) $\widehat{x} \in \mathcal{M}$ and $\|x - \widehat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\|$ if and only if $\widehat{x} \in \mathcal{M}$ and $(x - \widehat{x}) \in \mathcal{M}^{\perp}$ where $\mathcal{M}^{\perp}$ is the orthogonal complement to $\mathcal{M}$ in $\mathcal{H}$.

The element $\widehat{x}$ is called the orthogonal projection of x onto $\mathcal{M}$.

A.1. **Properties of Projection Mappings.** Let $\mathcal{H}$ be a Hilbert space and and let $P_{\mathcal{M}}$ be a projection mapping onto a closed subspace $\mathcal{M}$. Then

(i) each $x \in \mathcal{H}$ has a unique representation as a sum of of an element in $\mathcal{M}$ and an element in $\mathcal{M}^{\perp}$, i.e.

$$x = P_{\mathcal{M}}x + (I - P_{\mathcal{M}})x \tag{A.7}$$

(ii) $x \in \mathcal{M}$ if and only if $P_{\mathcal{M}}x = x$

(iii) $x \in \mathcal{M}^{\perp}$ if and only if $P_{\mathcal{M}}x = 0$

(iv) $\mathcal{M}_1 \subseteq \mathcal{M}_2$ if and only if $P_{\mathcal{M}2}P_{\mathcal{M}1}x = P_{\mathcal{M}1}$

(v) $\|x\|^2 = \|P_{\mathcal{M}}x\|^2 + \|(I - P_{\mathcal{M}})\, x\|^2$

Definition 5. *(The space $L^2(\Omega, F, P)$) We can define the space $L^2(\Omega, F, P)$ as the space consisting of all collections C of random variables X defined on the probability space (Ω, F, P) satisfying the condition*

$$EX^2 \int_{\Omega} X(\omega) P(d\omega) < \infty \tag{A.8}$$

and define the inner product of this space as

$$\langle X, Y \rangle = E(XY) \text{ for any } X, Y \in C \tag{A.9}$$