Kalman Filter and Maximum Likelihood Estimation of Linearized DSGE Models

Dr. Tai-kuang Ho*

*Associate Professor. Department of Quantitative Finance, National Tsing Hua University, No.

1 State space form and the Kalman filter

1.1 State space form

- x_t : the state variables

- z_t : the observable variables

- Transition equation

101, Section 2, Kuang-Fu Road, Hsinchu, Taiwan 30013, Tel: +886-3-571-5131, ext. 62136, Fax: +886-3-562-1823, E-mail: tkho@mx.nthu.edu.tw.

$$x_{t+1} = Fx_t + G\omega_{t+1}$$

$$\omega_{t+1} \sim \mathcal{N}(0, Q)$$

- Measurement equation

$$z_t = H_t' x_t + \upsilon_t$$

$$v_t \sim \mathcal{N}(0, R)$$

- We want to write the likelihood function of z_t.

1.2 Some useful properties of normal distribution

- Assume that

$$Z|w = \begin{bmatrix} X|w \\ Y|w \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix} \right)$$

- Then if follows

$$X|Y, w \sim \mathcal{N}\left(\bar{x} + \mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1}(Y - \bar{y}), \mathbf{\Sigma}_{xx} - \mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1}\mathbf{\Sigma}_{yx}\right)$$

1.3 Kalman filter

- $z^{t-1} \equiv \{z_{-1}, z_0, \ldots, z_{t-1}\}$
- $x_{t|t-1}$: the random variables x_t conditional on z^{t-1}, the history of the observable variables

- $x_{t|t-1} \equiv E\left(x_t|z^{t-1}\right)$
- $\Sigma_{t|t-1} \equiv E\left(\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)'|z^{t-1}\right)$
- $z_{t|t-1} \equiv E\left(z_t|z^{t-1}\right)$
- $\Omega_{t|t-1} \equiv E\left(\left(z_t - z_{t|t-1}\right)\left(z_t - z_{t|t-1}\right)'|z^{t-1}\right)$

$$\begin{aligned} z_{t|t-1} &\equiv E\left(z_t|z^{t-1}\right) = E\left(H_t' x_t + \upsilon_t|z^{t-1}\right) \\ &= H_t' x_{t|t-1} \end{aligned}$$

$$
\begin{aligned}
\Omega_{t|t-1} &\equiv E\left(\left(z_t - z_{t|t-1}\right)\left(z_t - z_{t|t-1}\right)' |z^{t-1}\right) \\
&= E\left(\left(H_t'\left(x_t - x_{t|t-1}\right) + \upsilon_t\right)\left(H_t'\left(x_t - x_{t|t-1}\right) + \upsilon_t\right)' |z^{t-1}\right) \\
&= E\left(H_t'\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' H_t|z^{t-1}\right) + E\left(\upsilon_t\upsilon_t'|z^{t-1}\right) \\
&= H_t'\Sigma_{t|t-1}H_t + R
\end{aligned}
$$

$$
\begin{aligned}
& E\left(\left(z_t - z_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' |z^{t-1}\right) \\
&= E\left(\left(H_t'\left(x_t - x_{t|t-1}\right) + \upsilon_t\right)\left(x_t - x_{t|t-1}\right)' |z^{t-1}\right) \\
&= H_t'\Sigma_{t|t-1}
\end{aligned}
$$

1.4 Kalman filter: first iteration

- $x_{0|-1}$

- $\Sigma_{0|-1}$ (symmetric matrix)

- Assume that

$$\left[\begin{array}{c} x_0 \\ z_0 \end{array} | z^{-1}\right] \sim \mathcal{N}\left(\begin{bmatrix} x_{0|-1} \\ H_0' x_{0|-1} \end{bmatrix}, \begin{bmatrix} \Sigma_{0|-1} & \Sigma_{0|-1} H_0 \\ H_0' \Sigma_{0|-1} & H_0' \Sigma_{0|-1} H_0 + R \end{bmatrix}\right)$$

- $z^{-1} \equiv \{z_{-1}\}$

- Then if follows

$$\left[x_0 | z^{-1}, z_0\right] \sim \mathcal{N}\left(x_{0|0}, \Sigma_{0|0}\right)$$

$$x_{0|0} = x_{0|-1} + \Sigma_{0|-1} H_0 \left(H_0' \Sigma_{0|-1} H_0 + R\right)^{-1} \left(z_0 - H_0' x_{0|-1}\right)$$

$$\Sigma_{0|0} = \Sigma_{0|-1} - \Sigma_{0|-1} H_0 \left(H_0' \Sigma_{0|-1} H_0 + R\right)^{-1} H_0' \Sigma_{0|-1}$$

- $z^0 \equiv \{z_{-1}, z_0\}$

- It follows that

$$
\begin{aligned}
x_{1|0} &= E\left(x_1 | z^0\right) = E\left(F x_0 + G\omega_1 | z^0\right) \\
&= F x_{0|0}
\end{aligned}
$$

$$
\begin{aligned}
\Sigma_{1|0} &= E\left(\left(x_1 - x_{1|0}\right)\left(x_1 - x_{1|0}\right)' | z^0\right) \\
&= E\left(\left(Fx_0 + G\omega_1 - x_{1|0}\right)\left(Fx_0 + G\omega_1 - x_{1|0}\right)' | z^0\right) \\
&= E\left(\left(Fx_0 + G\omega_1 - Fx_{0|0}\right)\left(Fx_0 + G\omega_1 - Fx_{0|0}\right)' | z^0\right) \\
&= E\left(F\left(x_0 - x_{0|0}\right)\left(x_0 - x_{0|0}\right)' F' | z^0\right) + E\left(G\left(\omega_1\right)\left(\omega_1\right)' G' | z^0\right) \\
&= F\Sigma_{0|0}F' + GQG'
\end{aligned}
$$

$$
\begin{aligned}
z_{1|0} &= E\left(z_1 | z^0\right) = E\left(H_1' x_1 + \upsilon_1 | z^0\right) \\
&= H_1' x_{1|0}
\end{aligned}
$$

$$
\begin{aligned}
\Omega_{1|0} &= E\left(\left(z_1 - z_{1|0}\right)\left(z_1 - z_{1|0}\right)' |z^0\right) \\
&= E\left(\left(H_1' x_1 + v_1 - H_1' x_{1|0}\right)\left(H_1' x_1 + v_1 - H_1' x_{1|0}\right)' |z^0\right) \\
&= E\left(H_1' \left(x_1 - x_{1|0}\right)\left(x_1 - x_{1|0}\right)' H_1 |z^0\right) + E\left(\left(v_1\right)\left(v_1\right)' |z^0\right) \\
&= H_1' \Sigma_{1|0} H_1 + R
\end{aligned}
$$

- We have moved as follows:

$$
\left(x_{0|-1}, \Sigma_{0|-1}\right) \to z_0 \to \left(x_{0|0}, \Sigma_{0|0}\right) \to \left(x_{1|0}, \Sigma_{1|0}, z_{1|0}, \Omega_{1|0}\right)
$$

1.5 Kalman filter: the algorithm

- $x_{t|t-1}$

- $\Sigma_{t|t-1}$

- z_t

$$x_{t|t} = x_{t|t-1} + \Sigma_{t|t-1} H_t \left(H_t' \Sigma_{t|t-1} H_t + R \right)^{-1} \left(z_t - H_t' x_{t|t-1} \right)$$

$$\boldsymbol{\Sigma}_{t|t} = \boldsymbol{\Sigma}_{t|t-1} - \boldsymbol{\Sigma}_{t|t-1} H_t \left(H_t' \boldsymbol{\Sigma}_{t|t-1} H_t + R \right)^{-1} H_t' \boldsymbol{\Sigma}_{t|t-1}$$

$$x_{t+1|t} = F x_{t|t}$$

$$\boldsymbol{\Sigma}_{t+1|t} = F \boldsymbol{\Sigma}_{t|t} F' + G Q G'$$

$$z_{t+1|t} = H_{t+1}' x_{t+1|t}$$

$$\Omega_{t+1|t} = H'_{t+1}\Sigma_{t+1|t}H_{t+1} + R$$

- We have moved as follows:

$$\left(x_{t|t-1}, \Sigma_{t|t-1}\right) \rightarrow z_t \rightarrow \left(x_{t|t}, \Sigma_{t|t}\right) \rightarrow \left(x_{t+1|t}, \Sigma_{t+1|t}, z_{t+1|t}, \Omega_{t+1|t}\right)$$

1.6 The likelihood function

- Prediction error decomposition of likelihood:

$$
\begin{aligned}
& \log \ell\left(z^{T} | F, G, H', Q, R\right) \\
= & \sum_{t=1}^{T} \log \ell\left(z_{t} | z^{t-1}, F, G, H', Q, R\right) \\
= & -\sum_{t=1}^{T}\left[\frac{N}{2} \log 2\pi + \frac{1}{2} \log \left|\Omega_{t|t-1}\right| + \frac{1}{2} \upsilon_{t}' \Omega_{t|t-1}^{-1} \upsilon_{t}\right]
\end{aligned}
$$

- $z_t | z^{t-1} \sim \mathcal{N}\left(z_{t|t-1}, \Omega_{t|t-1}\right)$

$$
N = \dim\left(z_t\right)
$$

$$\upsilon_t = z_t - z_{t|t-1} = z_t - H_t' x_{t|t-1}$$

$$\Omega_{t|t-1} = H_t' \Sigma_{t|t-1} H_t + R$$

1.7 Initial conditions for the Kalman filter

- All of the eigenvalues of F are inside the unit circle.
- GQG' and R are p.s.d. symmetric matrices.

- $H = \lim_{t \to \infty} H_t$

- We use the unconditional mean of x_t to initialize the Kalman filter.

- In other words, $x_{1|0} = \mathbf{0}$.

- We also use the unconditional variance of x_t to initialize the Kalman filter.

- In other words, we are setting $vec\left(\mathbf{\Sigma}_{1|0}\right) = (I - F \otimes F)^{-1} vec\left(GQG'\right)$.

$$x_{t+1} = Fx_t + G\omega_{t+1}$$

$$\boldsymbol{\Sigma} = F\boldsymbol{\Sigma}F' + GQG'$$

$$vec\left(\boldsymbol{\Sigma}\right) = vec\left(F\boldsymbol{\Sigma}F'\right) + vec\left(GQG'\right)$$

$$vec\left(\boldsymbol{\Sigma}\right) = \left(F \otimes F\right) vec\left(\boldsymbol{\Sigma}\right) + vec\left(GQG'\right)$$

$$vec\,(\boldsymbol{\Sigma}) = (I - F \otimes F)^{-1}\, vec\left(GQG'\right)$$

- An alternative is to choose the stationary value of the variance as the initial value for the variance.

$$\boldsymbol{\Sigma}_{t|t} = \boldsymbol{\Sigma}_{t|t-1} - \boldsymbol{\Sigma}_{t|t-1} H_t \left(H_t' \boldsymbol{\Sigma}_{t|t-1} H_t + R\right)^{-1} H_t' \boldsymbol{\Sigma}_{t|t-1}$$

$$\boldsymbol{\Sigma}_{t+1|t} = F \boldsymbol{\Sigma}_{t|t} F' + GQG'$$

$$\mathbf{\Sigma}_{t+1|t} = F\left[\mathbf{\Sigma}_{t|t-1} - \mathbf{\Sigma}_{t|t-1}H_t\left(H_t'\mathbf{\Sigma}_{t|t-1}H_t + R\right)^{-1}H_t'\mathbf{\Sigma}_{t|t-1}\right]F' + GQG'$$

$$\mathbf{\Sigma}_{t+1|t} \rightarrow \mathbf{\Sigma}$$

$$\mathbf{\Sigma} = F\left[\mathbf{\Sigma} - \mathbf{\Sigma}H\left(H'\mathbf{\Sigma}H + R\right)^{-1}H'\mathbf{\Sigma}\right]F' + GQG'$$

- $\mathbf{\Sigma}$ is the solution to the above algebraic Ricatti equation.

2 DSGE estimation

2.1 Express solution of DSGE models in state-space form

- We use the modified Paul Klein's code (solabHO.m) to solve the linearized system.

- Klein's code treats exogenous shocks as part of the state variables, and solves

$$A_0 \cdot \begin{bmatrix} k_{t+1} \\ u_{t+1} \end{bmatrix} = B_0 \cdot \begin{bmatrix} k_t \\ u_t \end{bmatrix}$$

- k_t : state variables plus exogenous shocks

- u_t : control variables

- Klein's code gives the results

$$k_{t+1} = p \cdot k_t \tag{1}$$

$$u_t = f \cdot k_t \tag{2}$$

- Assume that exogenous shocks follow:

$$Z_{t+1} = \Lambda Z_t + \tilde{\eta}\varepsilon_{t+1}$$

$$\varepsilon_{t+1} \sim \mathcal{N}\left(\mathbf{0}, V_1\right)$$

- You can follow Schmitt-Grohé and Uribe (2004, henceforth SGU) and assume that exogenous shocks follow:

$$Z_{t+1} = \Lambda Z_t + \tilde{\eta}\sigma\varepsilon_{t+1}$$

$$\varepsilon_{t+1} \sim \mathcal{N}(\mathbf{0}, I)$$

- We follow Ireland (2004) to transform the solution of DSGE models into a state-space form.

- First, the complete form of equation (1) is expressed as: (SGU, 2004, page 759):

$$k_{t+1} = p \cdot k_t + \begin{bmatrix} \mathbf{0} \\ \tilde{\eta} \end{bmatrix} \varepsilon_{t+1}$$

$$A \equiv p$$

$$B \equiv \begin{bmatrix} \mathbf{0} \\ \tilde{\eta} \end{bmatrix}$$

$$k_{t+1} = A \cdot k_t + B \cdot \varepsilon_{t+1} \quad (3)$$

- We use empirical and observable variables d_t to estimate the DSGE models.

- Second, equation (2) is related to observable variables d_t by:

$$d_t = C \cdot k_t$$

- If the observables d_t are contained in k_t, namely, they belong to the state variables of the models, then C is simple a selection matrix, such as:

$$C = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

- In this example, C matrix selects the first and last variables of vector k_t.

- In contrast, if the observables d_t are not contained in k_t, namely, they belong to the control variables of the models, then C is f pre-multiplied by a selection matrix (or is formed by simple taking out the corresponding rows of f), for example

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \cdot f$$

$$C = \begin{bmatrix} f_{2,:} \\ f_{3,:} \end{bmatrix}$$

- In this example, C matrix takes out the second and third rows of matrix f.

- In other words, the observables for estimation are the second and third control variables.

2.2 Stochastic singularity

- The number of exogenous shocks must be as least as many as the number of observable variables.

- Otherwise, there is the problem of stochastic singularity and estimation is impossible.

- Ingram, Kocherlakota, and Savin (1994) first notice the issue of stochastic singularity in the estimation of the real business cycle model with single technology shock.

- Ruge-Murcia (2007) provides a very good explanation to the issue of stochastic singularity.

- Here I draw on Ruge-Murcia (2007).

- Why stochastic singularity makes maximum likelihood estimation impossible?

- Stochastic singularity arises because DSGE models use a small number of structural shocks to generate predictions about a large number of observable variables.

- Therefore, the models predict that certain linear combinations of the observable variables should hold without noises, namely, be deterministic.

- Take Hansen's real business cycle with indivisible labor as an example.

- The model has only one technology shock.

- Observable variables: (y_t, n_t, c_t)

- State variables: (k_t)

- Exogenous structural shocks: (z_t)

- The solution of the endogenous observable variables is expressed as:

$$\underbrace{\begin{bmatrix} y_t \\ n_t \\ c_t \end{bmatrix}}_{s_t} = \underbrace{\begin{bmatrix} \phi_{yk} & \phi_{yz} \\ \phi_{nk} & \phi_{nz} \\ \phi_{ck} & \phi_{cz} \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} k_t \\ z_t \end{bmatrix}}_{x_t}$$

- Multiply out the decision rule:

$$\left\{ \begin{bmatrix} y_t = \phi_{yk} k_t + \phi_{yz} z_t \\ n_t = \phi_{nk} k_t + \phi_{nz} z_t \end{bmatrix} \right. \tag{A}$$

$$c_t = \phi_{ck} k_t + \phi_{cz} z_t \tag{B}$$

- Use (A) to solve for (k_t, z_t), and then substitute into (B) to obtain:

$$\left(\phi_{yk}\phi_{cz} - \phi_{yz}\phi_{ck}\right) n_t + (\phi_{nz}\phi_{ck} - \phi_{nk}\phi_{cz})\, y_t - \left(\phi_{nz}\phi_{yk} - \phi_{yz}\phi_{nk}\right) c_t = 0$$

- There exists a linear combination of observable variables that hold without noise.
- Under the models, the variance-covariance matrix of (y_t, n_t, c_t) is singular for any sample size and parameter values.

- $x_t \equiv (k_t, z_t)$
- $s_t \equiv (y_t, n_t, c_t)$
- $s^{t-1} \equiv \{s_{-1}, s_0, \ldots, s_{t-1}\}$
- $\Sigma_{t|t-1} \equiv E\left(\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' |s^{t-1}\right)$
- $\Omega_{t|t-1} \equiv E\left(\left(s_t - s_{t|t-1}\right)\left(s_t - s_{t|t-1}\right)' |s^{t-1}\right)$

$$s_t = Hx_t$$

$$s_t - s_{t|t-1} = H\left(x_t - x_{t|t-1}\right)$$

$$\Omega_{t|t-1} = H\Sigma_{t|t-1}H'$$

- We can follow the same procedure to show that elements of $\left(s_t - s_{t|t-1}\right)$ are not linear independent, since they are all proportional to the only technology shock.

- Therefore, the matrix $\Omega_{t|t-1} \equiv E\left(\left(s_t - s_{t|t-1}\right)\left(s_t - s_{t|t-1}\right)' | s^{t-1}\right)$ is singular.

- The inverse of $\Omega_{t|t-1}$ is not well defined and therefore likelihood cannot be evaluated.

2.3 Adding measurement errors

- To avoid the problem of stochastic singularity, we add measurement errors in the above equation:

$$d_t = C \cdot k_t + \upsilon_t \tag{4}$$

$$\upsilon_{t+1} = D \cdot \upsilon_t + \xi_{t+1} \tag{5}$$

$$\xi_{t+1} \sim \mathcal{N}\left(\mathbf{0}, V_2\right)$$

$$E\left(\varepsilon_{t+1} \cdot \xi'_{t+1}\right) = \mathbf{0}$$

- To address the issue of stochastic singularity, one can also estimate the model using at most as many observable variables as structural shocks, or to extend the model to permit additional structural shocks.

- While adding measurement errors preserves the original economic model, adding structural errors does not.

- Please refer to Ruge-Murcia (2007) and Tovar (2008) for further discussion.

- There are 3 ways to specify the measurement errors D and V_2.

- McGrattan et al. (1997), "An equilibrium model of the business cycle with household production and fiscal policy," *International Economic Review*: Add as many measurement errors as the number of observable variables.

- The matrices D and V_2 are restricted to be diagonal.

- For example,

$$D = \begin{bmatrix} \rho_1 & 0 & 0 \\ 0 & \rho_2 & 0 \\ 0 & 0 & \rho_3 \end{bmatrix}$$

$$V_2 = \begin{bmatrix} \sigma_{v1} & 0 & 0 \\ 0 & \sigma_{v2} & 0 \\ 0 & 0 & \sigma_{v3} \end{bmatrix}$$

- Canova (2007): Add measurement errors until the number of measurement errors plus the exogenous shocks are just equal to the number of observable variables.

- Matrix D is assumed to be a zero matrix.

- For example,

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$V_2 = \begin{bmatrix} \sigma_{v1} & 0 & 0 \\ 0 & \sigma_{v2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Ireland (2004): like McGrattan et al. (1997), but relax the restriction that matrices D and V_2 are diagonal.

- The eigenvalues of the matrix D are constrained to lie inside the unit circle, and the covariance matrix V_2 is constrained to the positive definite.

- For example,

$$D = \begin{bmatrix} \rho_1 & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_2 & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_3 \end{bmatrix}$$

$$V_2 = \begin{bmatrix} \sigma_{v1} & \sigma_{v12} & \sigma_{v13} \\ \sigma_{v12} & \sigma_{v2} & \sigma_{v23} \\ \sigma_{v13} & \sigma_{v23} & \sigma_{v3} \end{bmatrix}$$

- Here we follow McGrattan et al. (1997).

- Once this is done, we can proceed with the transformation, and employ the Kalman filter, and use the prediction error decomposition of likelihood to compute the likelihood value.

2.4 Use Kalman filter to compute likelihood function

- We define an augmented state vector:

$$x_t \equiv \begin{bmatrix} k_t \\ \upsilon_t \end{bmatrix}$$

$$\eta_{t+1} \equiv \begin{bmatrix} B \cdot \varepsilon_{t+1} \\ \xi_{t+1} \end{bmatrix}$$

- Equations (3), (4), and (5) can be rearranged as:

$$x_{t+1} = F x_t + \eta_{t+1} \tag{6}$$

$$d_t = Gx_t \tag{7}$$

$$F = \begin{bmatrix} A & \mathbf{0}_{nk\times nd} \\ \mathbf{0}_{nd\times nk} & D \end{bmatrix}$$

$$G = \begin{bmatrix} C & I_{nd\times nd} \end{bmatrix}$$

$$\eta_{t+\mathbf{1}} \sim \mathcal{N}\left(\mathbf{0}, Q\right)$$

$$Q \equiv E\left(\eta_{t+1}\eta'_{t+1}\right) = \begin{bmatrix} BV_1B' & \mathbf{0}_{nk\times nd} \\ \mathbf{0}_{nd\times nk} & V_2 \end{bmatrix}$$

- nk is the dimension of k_t, namely, the number of state variables plus exogenous shocks in the DSGE models.

- nd is the dimension of d_t, namely, the number of observable variables used in the estimation.

- Equations (6) and (7) constitute the state space form we employ in computing the likelihood function.

- The sequence of computation is as follows (we use the unconditional mean and variance of x_t to initialize the Kalman filter):

$$x_{1|0} = \mathbf{0}_{(nk+nd)\times 1}$$

$$vec\left(\boldsymbol{\Sigma}_{1|0}\right) = \left[I_{(nk+nd)^2\times(nk+nd)^2} - (F \otimes F)\right]^{-1} vec\,(Q)$$

$$\left(x_{t|t-1}, \boldsymbol{\Sigma}_{t|t-1}\right)$$

$$d_{t|t-1} \equiv E\left(d_t|d^{t-1}\right) = G \cdot x_{t|t-1}, \qquad d^{t-1} \equiv \{d_{-1}, d_0, \ldots, d_{t-1}\}$$

$$u_t = d_t - d_{t|t-1} = d_t - G \cdot x_{t|t-1}$$

$$\Omega_{t|t-1} \equiv \left[E\left(d_t - d_{t|t-1}\right)\left(d_t - d_{t|t-1}\right)' |d^{t-1}\right] = G \cdot \Sigma_{t|t-1} \cdot G'$$

$$K_{t|t-1} = F \cdot \Sigma_{t|t-1} \cdot G' \cdot \Omega_{t|t-1}^{-1}$$

$$x_{t+1|t} = F \cdot x_{t|t-1} + K_{t|t-1} \cdot u_t$$

$$\Sigma_{t+1|t} = Q + F \cdot \Sigma_{t|t-1} \cdot F' - K_{t|t-1} \cdot G \cdot \Sigma_{t|t-1} \cdot F'$$

$$\left(x_{t+1|t}, \Sigma_{t+1|t}\right)$$

$$\cdots$$

$$\log \ell = -\sum_{t=1}^{T} \left[\frac{nd}{2} \log 2\pi + \frac{1}{2} \log \left| \Omega_{t|t-1} \right| + \frac{1}{2} u_t' \Omega_{t|t-1}^{-1} u_t \right]$$

- Peter Ireland provides a very useful technical note to his 2004 paper (Ireland, 2003).

- I strongly recommend you to read it.

- The technical note of Ireland (2003) also explains how to compute impulse responses, variance decomposition, forecasting and to recover exogenous shocks.

- I use the transformation and formulation described here in the estimation of Monacelli (2005)'s small open economy model.

2.5 Extension

- Now instead of adding as many measurement errors as the number of observable variables, I will only add the number of measurement errors until the sum of number of exogenous shocks and measurement errors just are equal to the number of observables.

- The reason for doing so is that for a large-scale DSGE model, it is necessary to be sparse on parameters, otherwise convergence of the parameters estimation is a problem.

- I follow the notation of Ireland (2003).

$$s_{t+1} = As_t + B\varepsilon_{t+1}$$

$$d_t = Cs_t + P \cdot \upsilon_t$$

$$P_{nd\times(nd-nz)}$$

- P is a selection matrix.

- We assign measurement errors to the variables that are most likely to suffer from this measurement problem.

$$v_{t+1} = Dv_t + \xi_{t+1}$$

$$E\left(\varepsilon_{t+1}\varepsilon'_{t+1}\right) = V_1$$

$$E\left(\xi_{t+1}\xi_{t+1}'\right) = V_2$$

$$E\left(\varepsilon_{t+1}\xi_{t+1}'\right) = 0$$

$$x_t = \begin{bmatrix} s_t \\ v_t \end{bmatrix}$$

$$\eta_{t+1} = \begin{bmatrix} B\varepsilon_{t+1} \\ \xi_{t+1} \end{bmatrix}$$

$$x_{t+1} = Fx_t + \eta_{t+1}$$

$$d_t = Gx_t$$

$$F = \begin{bmatrix} A & 0_{nk \times (nd-nz)} \\ 0_{(nd-nz) \times nk} & D \end{bmatrix}$$

$$G = \begin{bmatrix} C & P \end{bmatrix}$$

$$\eta_{t+1} \sim \mathcal{N}(0, Q)$$

$$Q \equiv E\left(\eta_{t+1}\eta'_{t+1}\right) = \begin{bmatrix} BV_1B' & 0_{nk\times(nd-nz)} \\ 0_{(nd-nz)\times nk} & V_2 \end{bmatrix}$$

2.6 Anderson's transformation

- Hansen, Lars Peter, Ellen R. McGrattan, and Thomas J. Sargent, "Mechanics of forming and estimating dynamic linear economies," Federal Reserve Bank of Minneapolis, Research Department Staff Report 182, September 1994.

- Anderson, Evan W., Ellen R. McGrattan, Lars Peter Hansen, and Thomas J. Sargent (1996), "Mechanics of forming and estimating dynamic linear economies," in H.M. Amman, D.A. Kendrick, and J. Rust (eds.), *Handbook of Computational Economics*, Chapter 4, pp. 171-252, 1996.

- This is correct only when the vector white noise that drive the system is of the same dimension as that of vector of observable.

- A system (x_t, z_t) of the form:

$$x_{t+1} = A_o x_t + C\omega_{t+1} \tag{8}$$

$$z_t = Gx_t + \upsilon_t \tag{9}$$

$$\upsilon_t = D\upsilon_{t-1} + \eta_t$$

$$E\left(\omega_t \omega_t'\right) = I$$

$$E\left(\eta_t \eta_t'\right) = R$$

$$E\left(\omega_{t+1} \eta_s'\right) = 0, \quad \forall s, t$$

- Equation (8) is the transition equation.

- Equation (9) is the measurement equation.

- η_t : a martingale difference sequence

- ω_{t+1} is a martingale difference sequence with $E\left(\omega_t \omega_t'\right) = I$.

- D : a matrix whose eigenvalues are bounded in modulus by unity

- υ_t : a serially correlated measurement error process

- redefine

$$\bar{z}_t \equiv z_{t+1} - D z_t$$

$$\bar{G} = G A_o - D G$$

- state space system $(x_t, \bar{z}_t)$

$$x_{t+1} = A_o x_t + C \omega_{t+1}$$

$$\bar{z}_t = \bar{G}x_t + GC\omega_{t+1} + \eta_{t+1}$$

- The state noise remains $C\omega_{t+1}$
- The new measurement noise is $GC\omega_{t+1} + \eta_{t+1}$
- $K_t \equiv K_{t|t-1}$: the Kalman gain
- $\Sigma_t \equiv \Sigma_{t|t-1}$: the state covariance matrix

- It follows that

$$K_t = \left(CC'G' + A_o\mathbf{\Sigma}_t\bar{G}'\right)\Omega_t^{-1}$$

$$\Omega_t = \bar{G}\mathbf{\Sigma}_t\bar{G}' + R + GCC'G'$$

$$\begin{aligned}\mathbf{\Sigma}_{t+1} &= A_0\mathbf{\Sigma}_t A_0' + CC' - \left(CC'G' + A_o\mathbf{\Sigma}_t\bar{G}'\right)\Omega_t^{-1}\left(\bar{G}\mathbf{\Sigma}_t A_o' + GCC'\right) \\ &= A_0\mathbf{\Sigma}_t A_0' + CC' - K_t\left(\bar{G}\mathbf{\Sigma}_t A_o' + GCC'\right)\end{aligned}$$

- The state space system can be expressed in an innovation representation:

$$u_t = \bar{z}_t - \bar{G}x_{t|t-1}$$

$$x_{t+1|t} = A_o x_{t|t-1} + K_t u_t$$

- I use this transformation and formulation in the estimation of Chari, Kehoe and McGrattan (2007)'s prototype model for business cycle accounting.

- Christensen, Hurn, and Lindsay (2008) provide useful tips for numerical optimization.

References

[1] Canova, Fabio (2007), *Methods for Applied Macroeconomic Research*, Chapter 6: Likelihood Methods.

[2] Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan (2007), "Business cycle accounting," *Econometrica*, 75 (3), pp. 781-836.

[3] Christensen, T. M., A. S. Hurn, and K. A. Lindsay (2008), "The Devil is in the Detail: Hints for Practical Optimization"

[4] Ingram, Beth Fisher, Narayana R. Kocherlakota, and N. E. Savin (1994), "Explaining business cycles: a multiple-shock approach," *Journal of Monetary Economics*, 34, pp. 415-428.

[5] Ireland, Peter N. (2004), "A method for taking models to the data," *Journal of Economic Dynamics and Control*, 28, pp. 1205-1226.

[6] Ireland, Peter N. (2003), Technical Note on "A method for taking models to the data," available from Peter Ireland's website.

[7] McGrattan, Ellen, Richard Rogerson, and Randall Wright (1995), "An equilibrium model of the business cycle with household production and fiscal policy," Federal Reserve Bank of Minneapolis Research Department Staff Report 191.

[8] McGrattan, Ellen, Richard Rogerson, and Randall Wright (1997), "An equilibrium model of the business cycle with household production and fiscal policy," *International Economic Review*, 38 (2), pp. 267-90.

[9] Ruge-Murcia, Francisco J. (2007), "Methods to estimate dynamic stochastic general equilibrium models," *Journal of Economic Dynamics and Control*, 31, pp. 2599-2636.

[10] Tovar, Camilo E. (2008), "DSGE models and central banks," BIS Working Papers No. 258.