Linear DSGE Models and the Kalman Filter

Ed Herbst

November 10, 2020

Background

- *Textbook treatments:* Woodford (2003), Galí (2008)
- *Key empirical papers*: Ireland (2004), Christiano et al. (2005), Smets and Wouters (2007), An and Schorfheide (2007),
- *Frequentist estimation:* Harvey (1991), Hamilton (1994),
- *Bayesian estimation:* Herbst and Schorfheide (2015)

Small-Scale DSGE Model

- Intermediate and final goods producers
- Households
- Monetary and fiscal policy
- Exogenous processes
- Equilibrium Relationships

Final Goods Producers

- Perfectly competitive firms combine a continuum of intermediate goods:

$$Y_t = \left(\int_0^1 Y_t(j)^{1-\nu} dj \right)^{\frac{1}{1-\nu}}.$$

- Firms take input prices $P_t(j)$ and output prices P_t as given; maximize profits

$$\Pi_t = P_t \left(\int_0^1 Y_t(j)^{1-\nu} dj \right)^{\frac{1}{1-\nu}} - \int_0^1 P_t(j) Y_t(j) dj.$$

- Demand for intermediate good j:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\nu} Y_t.$$

- Zero-profit condition implies

$$P_t = \left(\int_0^1 P_t(j)^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}.$$

Intermediate Goods Producers

- Intermediate good j is produced by a monopolist according to:
$$Y_t(j) = A_t N_t(j).$$
- Nominal price stickiness via quadratic price adjustment costs
$$AC_t(j) = \frac{\phi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - \pi\right)^2 Y_t(j).$$
- Firm j chooses its labor input $N_t(j)$ and the price $P_t(j)$ to maximize the present value of future profits:
$$\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s Q_{t+s|t}\left(\frac{P_{t+s}(j)}{P_{t+s}}Y_{t+s}(j) - W_{t+s}N_{t+s}(j) - AC_{t+s}(j)\right)\right].$$

Households

- Household derives disutility from hours worked H_t and maximizes

$$\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s\left(\frac{(C_{t+s}/A_{t+s})^{1-\tau}-1}{1-\tau}\right.\right.$$
$$\left.\left.+\chi_M \ln\left(\frac{M_{t+s}}{P_{t+s}}\right)-\chi_H H_{t+s}\right)\right].$$

- Budget constraint:

$$\begin{aligned} & P_t C_t + B_t + M_t + T_t \\ & \quad = \quad P_t W_t H_t + R_{t-1}B_{t-1} + M_{t-1} + P_t D_t + P_t SC_t. \end{aligned}$$

Monetary and Fiscal Policy

- Central bank adjusts money supply to attain desired interest rate.
- Monetary policy rule:

$$
\begin{aligned}
R_t &= R_t^{*,1-\rho_R} R_{t-1}^{\rho_R} e^{\epsilon_{R,t}} \\
R_t^* &= r\pi^* \left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{Y_t^*}\right)^{\psi_2}
\end{aligned}
$$

- Fiscal authority consumes fraction of aggregate output: $G_t = \zeta_t Y_t$.
- Government budget constraint:

$$
P_t G_t + R_{t-1} B_{t-1} + M_{t-1} = T_t + B_t + M_t.
$$

Exogenous Processes

- Technology:

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t, \quad \ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t}.$$

- Government spending / aggregate demand: define $g_t = 1/(1-\zeta_t)$; assume

$$\ln g_t = (1-\rho_g)\ln g + \rho_g \ln g_{t-1} + \epsilon_{g,t}.$$

- Monetary policy shock $\epsilon_{R,t}$ is assumed to be serially uncorrelated.

Equilibrium Conditions

- Consider the symmetric equilibrium in which all intermediate goods producing firms make identical choices; omit j subscript.
- Market clearing:

$$Y_t = C_t + G_t + AC_t \quad \text{and} \quad H_t = N_t.$$

- Complete markets:

$$Q_{t+s|t} = (C_{t+s}/C_t)^{-\tau}(A_t/A_{t+s})^{1-\tau}.$$

- Consumption Euler equation and New Keynesian Phillips curve:

$$\begin{aligned} 1 &= \beta \mathbb{E}_t \left[\left(\frac{C_{t+1}/A_{t+1}}{C_t/A_t} \right)^{-\tau} \frac{A_t}{A_{t+1}} \frac{R_t}{\pi_{t+1}} \right] \\ 1 &= \phi(\pi_t - \pi) \left[\left(1 - \frac{1}{2\nu} \right) \pi_t + \frac{\pi}{2\nu} \right] \\ &\quad -\phi\beta\mathbb{E}_t \left[\left(\frac{C_{t+1}/A_{t+1}}{C_t/A_t} \right)^{-\tau} \frac{Y_{t+1}/A_{t+1}}{Y_t/A_t} (\pi_{t+1} - \pi)\pi_{t+1} \right] \end{aligned}$$

Equilibrium Conditions – Continued

- In the absence of nominal rigidities ($\phi = 0$) aggregate output is given by

$$Y_t^* = (1 - \nu)^{1/\tau} A_t g_t,$$

which is the target level of output that appears in the monetary policy rule.

Steady State

- Set $\epsilon_{R,t}$, $\epsilon_{g,t}$, and $\epsilon_{z,t}$ to zero at all times.
- Because technology $\ln A_t$ evolves according to a random walk with drift $\ln \gamma$, consumption and output need to be detrended for a steady state to exist.
- Let
$$c_t = C_t/A_t, \quad y_t = Y_t/A_t, \quad y_t^* = Y_t^*/A_t.$$
- Steady state is given by:
$$\begin{aligned} \pi &= \pi^*, \quad r = \frac{\gamma}{\beta}, \quad R = r\pi^*, \\ c &= (1-\nu)^{1/\tau}, \quad y = gc = y^*. \end{aligned}$$

Solving DSGE Models

- Derive nonlinear equilibrium conditions:
  - System of nonlinear expectational difference equations;
  - transversality conditions.
- Find solution(s) of system of expectational difference methods:
  - Global (nonlinear) approximation
  - Local approximation near steady state
- We will focus on log-linear approximations around the steady state.
- More detail in: Fernandez-Villaverde et al. (2016): “Solution and Estimation Methods for DSGE Models.”

What is a Local Approximation?

- In a nutshell… consider the backward-looking model

$$y_t = f(y_{t-1}, \sigma\epsilon_t).$$

- Guess that the solution is of the form

$$y_t = y_t^{(0)} + \sigma y_t^{(1)} + o(\sigma).$$

- Steady state:

$$y_t^{(0)} = y^{(0)} = f(y^{(0)}, 0)$$

- Suppose $y^{(0)} = 0$. Expand $f(\cdot)$ around $\sigma = 0$:

$$f(y_{t-1}, \sigma\epsilon_t) = f_y y_{t-1} + f_\epsilon \sigma\epsilon_t + o(|y_{t-1}|) + o(\sigma)$$

- Now plug-in conjectured solution:

$$\sigma y_t^{(1)} = f_y \sigma y_{t-1}^{(1)} + f_\epsilon \sigma\epsilon_t + o(\sigma)$$

- Deduce that $y_t^{(1)} = f_y y_{t-1}^{(1)} + f_\epsilon \epsilon_t$

What is a Log-Linear Approximation?

- Consider a Cobb-Douglas production function: $Y_t = A_t K_t^{\alpha} N_t^{1-\alpha}$.
- Linearization around Y_*, A_*, K_*, N_*:

$$\begin{aligned} Y_t - Y_* &\approx K_*^{\alpha} N_*^{1-\alpha}(A_t - A_*) + \alpha A_* K_*^{\alpha-1} N_*^{1-\alpha}(K_t - K_*) \\ &\quad +(1-\alpha) A_* K_*^{\alpha} N_*^{-\alpha}(N_t - N_*) \end{aligned}$$

- Log-linearization: Let $f(x) = f(e^v)$ and linearize with respect to v:

$$f(e^v) \approx f(e^{v_*}) + e^{v_*} f'(e^{v_*})(v - v_*).$$

Thus:

$$f(x) \approx f(x_*) + x_* f'(x_*)(\ln x / x_*) = f(x_*) + f'(x_*)\tilde{x}$$

- Cobb-Douglas production function (here relationship is exact):

$$\tilde{Y}_t = \tilde{A}_t + \alpha \tilde{K}_t + (1-\alpha)\tilde{N}_t$$

Loglinearization of New Keynesian Model

- Consumption Euler equation:

$$\hat{y}_t = \mathbb{E}_t[\hat{y}_{t+1}] - \frac{1}{\tau}\left(\hat{R}_t - \mathbb{E}_t[\hat{\pi}_{t+1}] - \mathbb{E}_t[\hat{z}_{t+1}]\right) + \hat{g}_t - \mathbb{E}_t[\hat{g}_{t+1}]$$

- New Keynesian Phillips curve:

$$\hat{\pi}_t = \beta\mathbb{E}_t[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{g}_t),$$

where

$$\kappa = \tau\frac{1-\nu}{\nu\pi^2\phi}$$

- Monetary policy rule:

$$\hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)\psi_1\hat{\pi}_t + (1-\rho_R)\psi_2\left(\hat{y}_t - \hat{g}_t\right) + \epsilon_{R,t}$$

Canonical Linear Rational Expectations System

- Define
$$x_t = [\hat{y}_t, \hat{\pi}_t, \hat{R}_t, \epsilon_{R,t}, \hat{g}_t, \hat{z}_t]'.$$
- Augment x_t by $\mathbb{E}_t[\hat{y}_{t+1}]$ and $\mathbb{E}_t[\hat{\pi}_{t+1}]$.
- Define
$$s_t = \left[x_t', \mathbb{E}_t[\hat{y}_{t+1}], \mathbb{E}_t[\hat{\pi}_{t+1}]\right]'.$$
- Define rational expectations forecast errors forecast errors for inflation and output. Let
$$\eta_{y,t} = \hat{y}_t - \mathbb{E}_{t-1}[\hat{y}_t], \quad \eta_{\pi,t} = \hat{\pi}_t - \mathbb{E}_{t-1}[\hat{\pi}_t].$$
- Write system in canonical form Sims (2002):
$$\Gamma_0 s_t = \Gamma_1 s_{t-1} + \Psi \epsilon_t + \Pi \eta_t.$$

How Can One Solve Linear Rational Expectations Systems? A Simple Example

- Consider

$$y_t = \frac{1}{\theta}\mathbb{E}_t[y_{t+1}] + \epsilon_t, \tag{1}$$

  where $\epsilon_t \sim iid(0,1)$ and $\theta \in \Theta = [0,2]$.

- Introduce conditional expectation $\xi_t = \mathbb{E}_t[y_{t+1}]$ and forecast error $\eta_t = y_t - \xi_{t-1}$.

- Thus,

$$\xi_t = \theta\xi_{t-1} - \theta\epsilon_t + \theta\eta_t. \tag{2}$$

A Simple Example

- Determinacy: $\theta > 1$. Then only stable solution:

$$\xi_t = 0, \quad \eta_t = \epsilon_t, \quad y_t = \epsilon_t \tag{3}$$

- Indeterminacy: $\theta \leq 1$ the stability requirement imposes no restrictions on forecast error:

$$\eta_t = \widetilde{M}\epsilon_t + \zeta_t. \tag{4}$$

- For simplicity assume now $\zeta_t = 0$. Then

$$y_t - \theta y_{t-1} = \widetilde{M}\epsilon_t - \theta\epsilon_{t-1}. \tag{5}$$

- General solution methods for LREs: Blanchard and Kahn (1980), King and Watson (1998), Uhlig (1999), Anderson (2000), Klein (2000), Christiano (2002), Sims (2002).

Solving a More General System

- Canonical form:

$$\Gamma_0(\theta)s_t = \Gamma_1(\theta)s_{t-1} + \Psi(\theta)\epsilon_t + \Pi(\theta)\eta_t, \tag{6}$$

- The system can be rewritten as

$$s_t = \Gamma_1^*(\theta)s_{t-1} + \Psi^*(\theta)\epsilon_t + \Pi^*(\theta)\eta_t. \tag{7}$$

- Replace Γ_1^* by $J\Lambda J^{-1}$ and define $w_t = J^{-1}s_t$.
- To deal with repeated eigenvalues and non-singular Γ_0 we use Generalized Complex Schur Decomposition (QZ) in practice.
- Let the i'th element of w_t be $w_{i,t}$ and denote the i'th row of $J^{-1}\Pi^*$ and $J^{-1}\Psi^*$ by $[J^{-1}\Pi^*]_{i.}$ and $[J^{-1}\Psi^*]_{i.}$, respectively.

Solving a More General System

- Rewrite model:

$$w_{i,t} = \lambda_i w_{i,t-1} + [J^{-1}\Psi^*]_{i.}\epsilon_t + [J^{-1}\Pi^*]_{i.}\eta_t. \tag{8}$$

- Define the set of stable AR(1) processes as

$$I_s(\theta) = \left\{ i \in \{1, \dots n\} \middle| |\lambda_i(\theta)| \leq 1 \right\} \tag{9}$$

- Let $I_x(\theta)$ be its complement. Let Ψ_x^J and Π_x^J be the matrices composed of the row vectors $[J^{-1}\Psi^*]_{i.}$ and $[J^{-1}\Pi^*]_{i.}$ that correspond to unstable eigenvalues, i.e., $i \in I_x(\theta)$.
- Stability condition:

$$\Psi_x^J \epsilon_t + \Pi_x^J \eta_t = 0 \tag{10}$$

for all t.

Solving a More General System

- Solving for η_t. Define

$$
\begin{aligned}
\Pi_x^J &= \begin{bmatrix} U_{.1} & U_{.2} \end{bmatrix} \begin{bmatrix} D_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V'_{.1} \\ V'_{.2} \end{bmatrix} \qquad (11) \\
&= \underbrace{U}_{m \times m} \underbrace{D}_{m \times k} \underbrace{V'}_{k \times k} \\
&= \underbrace{U_{.1}}_{m \times r} \underbrace{D_{11}}_{r \times r} \underbrace{V'_{.1}}_{r \times k}.
\end{aligned}
$$

- If there exists a solution to Eq. (10) that expresses the forecast errors as function of the fundamental shocks ϵ_t and sunspot shocks ζ_t, it is of the form

$$
\begin{aligned}
\eta_t &= \eta_1 \epsilon_t + \eta_2 \zeta_t \qquad (12) \\
&= (-V_{.1} D_{11}^{-1} U'_{.1} \Psi_x^J + V_{.2} \widetilde{M}) \epsilon_t + V_{.2} M_\zeta \zeta_t,
\end{aligned}
$$

where $\widetilde{M}$ is an $(k-r) \times l$ matrix, M_ζ is a $(k-r) \times p$ matrix, and the dimension of $V_{.2}$ is $k \times (k-r)$. The solution is unique if $k = r$ and $V_{.2}$ is zero.

Proposition

If there exists a solution to Eq. (10) that expresses the forecast errors as function of the fundamental shocks ϵ_t and sunspot shocks ζ_t, it is of the form

$$\begin{aligned} \eta_t &= \eta_1 \epsilon_t + \eta_2 \zeta_t \\ &= (-V_{.1} D_{11}^{-1} U_{.1}' \Psi_x^J + V_{.2} \widetilde{M}) \epsilon_t + V_{.2} M_\zeta \zeta_t, \end{aligned} \tag{13}$$

where $\widetilde{M}$ is an $(k-r) \times l$ matrix, M_ζ is a $(k-r) \times p$ matrix, and the dimension of $V_{.2}$ is $k \times (k-r)$. The solution is unique if $k = r$ and $V_{.2}$ is zero.

At the End of the Day…

- We obtain a transition equation for the vector s_t:

$$s_t = T(\theta)s_{t-1} + R(\theta)\epsilon_t.$$

- The coefficient matrices $T(\theta)$ and $R(\theta)$ are functions of the parameters of the DSGE model.

Measurement Equation

- Relate model variables s_t to observables y_t.
- In NK model:

$$
\begin{aligned}
YGR_t &= \gamma^{(Q)} + 100(\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t) \\
INFL_t &= \pi^{(A)} + 400\hat{\pi}_t \\
INT_t &= \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} + 400\hat{R}_t.
\end{aligned}
$$

where

$$
\gamma = 1 + \frac{\gamma^{(Q)}}{100}, \quad \beta = \frac{1}{1 + r^{(A)}/400}, \quad \pi = 1 + \frac{\pi^{(A)}}{400}.
$$

- More generically:

$$
y_t = D(\theta) + Z(\theta)s_t + \underbrace{\eta_t}_{\text{optional}} .
$$

The state and measurement equations define a *State Space Model*.

State Space Models

- State space models form a very general class of models that encompass many of the specifications that we encountered earlier.
- ARMA models and linearized DSGE models can be written in state space form.

A state space model consists of

- a measurement equation that relates an *unobservable* state vector s_t to the *observables* y_t,
- a transition equation that describes the evolution of the state vector s_t.

Measurement Equation

The measurement equation is of the form

$$y_t = D_{t|t-1} + Z_{t|t-1} s_t + \eta_t, \quad t = 1, \ldots, T \tag{14}$$

where y_t is a $n_y \times 1$ vector of observables, s_t is a $n_s \times 1$ vector of state variables, $Z_{t|t-1}$ is an $n_y \times n_s$ vector, $D_{t|t-1}$ is a $n_y \times 1$ vector, and η_t are innovations (or often "measurement errors") with mean zero and $\mathbb{E}_{t-1}[\eta_t \eta_t'] = H_{t|t-1}$.

- The matrices $Z_{t|t-1}$, $D_{t|t-1}$, and $H_{t|t-1}$ are in many applications constant.
- However, it is sufficient that they are predetermined at $t-1$. They could be functions of $y_{t-1}, y_{t-2}, \ldots$.
- To simplify the notation, we will denote them by Z_t, D_t, and H_t, respectively.

Transition Equation

The transition equation is of the form

$$s_t = C_{t|t-1} + T_{t|t-1} s_{t-1} + R_{t|t-1} \epsilon_t \tag{15}$$

where R_t is $n_s \times n_\epsilon$, and ϵ_t is a $n_\epsilon \times 1$ vector of innovations with mean zero and variance $\mathbb{E}_{t|t-1}[\epsilon_t \epsilon_t'] = Q_{t|t-1}$.

- The assumption that s_t evolves according to an VAR(1) process is not very restrictive, since it could be the companion form to a higher order VAR process.
- It is furthermore assumed that (i) expectation and variance of the initial state vector are given by $\mathbb{E}[s_0] = A_0$ and $var[s_0] = P_0$;
- ϵ_t and η_t are uncorrelated with each other in all time periods , and uncorrelated with the initial state. [not really necessary]

Adding it all up

If the system matrices $Z_t, D_t, H_t, T_t, C_t, R_t, Q_t$ are non-stochastic and predetermined, then the system is linear and y_t can be expressed as a function of present and past ϵ_t's and η_t's.

1. calculate predictions $y_t|Y^{t-1}$, where $Y^{t-1} = [y_{t-1}, \ldots, y_1]$,
2. obtain a likelihood function

$$p(Y^T|\{Z_t, D_t, H_t, T_t, C_t, R_t, Q_t\})$$

3. back out a sequence

$$\{p(s_t|Y^t, \{Z_t, D_t, H_t, T_t, c_t, R_t, Q_t\})\}$$

The algorithm is called the *Kalman Filter* and was originally adopted from the engineering literature.

A Useful Lemma

Let $(x', y')'$ *be jointly normal with*

$$\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Then the $pdf(x|y)$ is multivariate normal with

$$\begin{aligned} \mu_{x|y} &= \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y) \\ \Sigma_{xx|y} &= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \quad \Box. \end{aligned}$$

(16)

A Bayesian Interpretation to the Kalman Filter

- Although the idea of the algorithm is based on linear projections, it has a very straightforward Bayesian interpretation.
- We will assume that the conditional distributions of s_t and y_t given time $t-1$ information are Gaussian.
- Since the system is linear, all the conditional and marginal distributions that we calculate when we move from period $t-1$ to period t will also be Gaussian.
- Since the state vector s_t is unobservable, it is natural in Bayesian framework to regard it as a random vector.

Note: The subsequent analysis is conditional on the system matrices $Z_t, D_t, H_t, T_t, C_t, R_t, Q_t$. For notational convenience we will, however, drop the system matrices from the conditioning set.

The calculations will be based on the following conditional distribution, represented by densities:

1. Initialization: $p(s_{t-1}|Y^{t-1})$
2. Forecasting:

$$
\begin{aligned}
p(s_t|Y^{t-1}) &= \int p(s_t|s_{t-1}, Y^{t-1})p(s_{t-1}|Y^{t-1})ds_{t-1} \\
p(y_t|Y^{t-1}) &= \int p(y_t|s_t, Y^{t-1})p(s_t|Y^{t-1})ds_t
\end{aligned}
$$

3. Updating:

$$
p(s_t|Y^t) = \frac{p(y_t|s_t, Y^{t-1})p(s_t|Y_{t-1})}{p(y_t|Y^{t-1})}
$$

- The integrals look troublesome.
- However, since the state space model is linear, and the distribution of the innovations u_t and η_t are Gaussian $\Longrightarrow$ everything is Gaussian!
- Hence, we only have to keep track of conditional means and variances.

Initialization

- In period zero, we will start with a prior distribution for the initial state s_0.

- This prior is of the form $s_0 \sim \mathcal{N}(A_0, P_0)$.

- If the system matrices imply that the state vector has a stationary distribution, we could choose A_0 and P_0 to be the mean and variance of this stationary distribution.

Forecasting

- At $(t-1)^+$, that is, after observing y_{t-1}, the belief about the state vector has the form $s_{t-1}|Y^{t-1} \sim (A_{t-1}, P_{t-1})$.

- Thus, the "posterior" from period $t-1$ turns into a prior for $(t-1)^+$.

Since s_{t-1} and η_t are independent multivariate normal random variables, it follows that

$$s_t|Y^{t-1} \sim \mathcal{N}(\hat{s}_{t|t-1}, P_{t|t-1}) \tag{17}$$

where

$$\begin{aligned} \hat{s}_{t|t-1} &= T_t A_{t-1} + c_t \\ P_{t|t-1} &= T_t P_{t-1} T_t' + R_t Q_t R_t' \end{aligned}$$

Forecasting y_t

The conditional distribution of $y_t | s_t, Y^{t-1}$ is of the form

$$y_t | s_t, Y^{t-1} \sim \mathcal{N}(Z_t s_t + D_t, H_t) \tag{18}$$

Since $s_t | Y^{t-1} \sim \mathcal{N}(\hat{s}_{t|t-1}, P_{t|t-1})$, we can deduce that the marginal distribution of y_t conditional on Y^{t-1} is of the form

$$y_t | Y_{t-1} \sim \mathcal{N}(\hat{y}_{t|t-1}, F_{t|t-1}) \tag{19}$$

where

$$\begin{aligned} \hat{y}_{t|t-1} &= Z_t \hat{s}_{t|t-1} + d_t \\ F_{t|t-1} &= Z_t P_{t|t-1} Z_t' + H_t \end{aligned}$$

Updating

To obtain the posterior distribution of $s_t | y_t, Y^{t-1}$ note that

$$\begin{aligned} s_t &= \hat{s}_{t|t-1} + (s_t - \hat{s}_{t|t-1}) && (20) \\ y_t &= Z_t \hat{s}_{t|t-1} + D_t + Z_t (s_t - \hat{s}_{t|t-1}) + u_t && (21) \end{aligned}$$

and the joint distribution of s_t and y_t is given by

$$\left[\begin{array}{c} s_t \\ x_t \end{array} \right] \Big| Y^{t-1} \sim \mathcal{N} \left(\left[\begin{array}{c} \hat{s}_{t|t-1} \\ \hat{y}_{t|t-1} \end{array} \right], \left[\begin{array}{cc} P_{t|t-1} & P_{t|t-1} Z_t' \\ Z_t P_{t|t-1}' & F_{t|t-1} \end{array} \right] \right) \quad (22)$$

$$s_t | y_t, Y^{t-1} \sim \mathcal{N}(A_t, P_t) \quad (23)$$

where

$$\begin{aligned} A_t &= \hat{s}_{t|t-1} + P_{t|t-1} Z_t' F_{t|t-1}^{-1} (y_t - Z_t \hat{s}_{t|t-1} - d_t) \\ P_t &= P_{t|t-1} - P_{t|t-1} Z_t' F_{t|t-1}^{-1} Z_t P_{t|t-1} \end{aligned}$$

The conditional mean and variance $\hat{y}_{t|t-1}$ and $F_{t|t-1}$ were given above. This completes one iteration of the algorithm. The posterior $s_t | Y^t$ will serve as prior for the next iteration. □

Likelihood Function

We can define the one-step ahead forecast error

$$\nu_t = y_t - \hat{y}_{t|t-1} = Z_t(s_t - \hat{s}_{t|t-1}) + u_t \tag{24}$$

The likelihood function is given by

$$\begin{aligned} p(Y^T|\text{parameters}) &= \prod_{t=1}^{T} p(y_t|Y^{t-1}, \text{parameters}) \\ &= (2\pi)^{-nT/2} \left(\prod_{t=1}^{T} |F_{t|t-1}| \right)^{-1/2} \\ &\quad \times \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \nu_t F_{t|t-1} \nu_t' \right\} \end{aligned} \tag{25}$$

This representation of the likelihood function is often called prediction error form, because it is based on the recursive prediction one-step ahead prediction errors ν_t. $\square$

Multistep Forecasting

The Kalman Filter can also be used to obtain multi-step ahead forecasts. For simplicity, suppose that the system matrices are constant over time. Since

$$s_{t+h-1|t-1} = T^h s_{t-1} + \sum_{s=0}^{h-1} T^s c + \sum_{s=0}^{h-1} T^s R \eta_t \tag{26}$$

it follows that

$$\begin{aligned}
\hat{s}_{t+h-1|t-1} &= [s_{t+h-1|t-1}|Y^{t-1}] = T^h A_{t-1} + \sum_{s=0}^{h-1} T^s c \\
P_{t+h-1|t-1} &= var[s_{t+h-1|t-1}|Y^{t-1}] = T^h P_{t-1} T^h + \sum_{s=0}^{h-1} T^s RQR' T^{s'}
\end{aligned}$$

which leads to

$$y_{t+h-1}|Y_{t-1} \sim \mathcal{N}(\hat{y}_{t+h-1|t-1}, F_{t+h-1|t-1}) \tag{27}$$

where

$$\hat{y}_{t+h-1|t-1} = Z\hat{s}_{t+h-1|t-1} + d$$

Some Discussion

- *Initialization*. If s_t is covariance stationary, can set (A_0, P_0), based on invariant distribution, otherwise, P_0 is typically extremely large, like $1000 \times I_n$

- *Kalman Gain*.
$$K_t = P_{t|t-1} Z F_{t|t-1}^{-1},$$
is an $n_s \times n_y$ matrix that maps the "surprises" (forecast errors) in the observed data to changes in our beliefs about the underlying unobserved states.

- *Missing data.* KF easily handles missing data through a change in the observation equation.

- *Kalman Smoother* delivers distributions, $\{s_t | Y^T\}_{t=1}^{T}$.

Relationship between State Space and VAR Models?

Question: Can we write the state space model as a VAR?
Assume the system matrices are time invariant,
$C = D = H = 0$, and $n_y = n_\epsilon$.
Let's write the model slightly differently:

$$
\begin{aligned}
s_t &= \mathbf{A} s_{t-1} + \mathbf{B}\epsilon_t && (28) \\
y_t &= \mathbf{C} s_{t-1} + \mathbf{D}\epsilon_t && (29)
\end{aligned}
$$

$\mathbf{A} = T, \mathbf{B} = R, \mathbf{C} = ZT, \mathbf{D} = ZR$
This means that

$$\epsilon_t = (\mathbf{D})^{-1}(y_t - \mathbf{C} s_{t-1}).$$

Using the state equation

$$s_t = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) s_{t-1} + \mathbf{B}\mathbf{D}^{-1} y_t.$$

Solving backwards,

$$s_t = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{t-1} s_0 + \sum_{j=0} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{j-1} \mathbf{B}\mathbf{D}^{-1} y_{t-j}$$

If eigenvalues of $(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})$ are less than one in modulus, then

$$lim_{t\to\infty} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^t \to 0$$

And we can write the states as a combination of the history of observations. So

$$y_t \approx \mathbf{C} \sum_{j=0}^{\infty} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{j-1} \mathbf{B}\mathbf{D}^{-1} y_{t-1-j} + \mathbf{D}\epsilon_t.$$

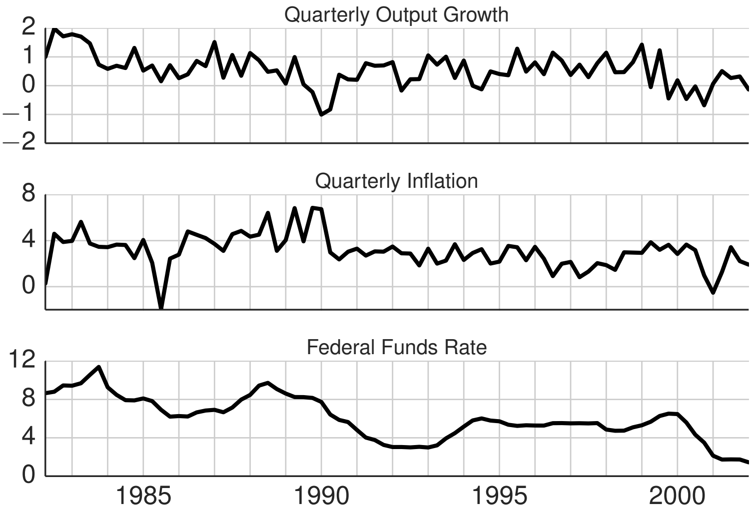
We have a VAR(∞) representation for y_t whose innovations coincide with the structural shocks of our DSGE model!

The condition that *eigenvalues of* $(\mathbf{A} - \mathbf{B}(\mathbf{D})^{-1}\mathbf{C})$ *are less than one in modulus* is known as the Poor Man's Invertibility Condition, see Fernndez-Villaverde et al. (2007).

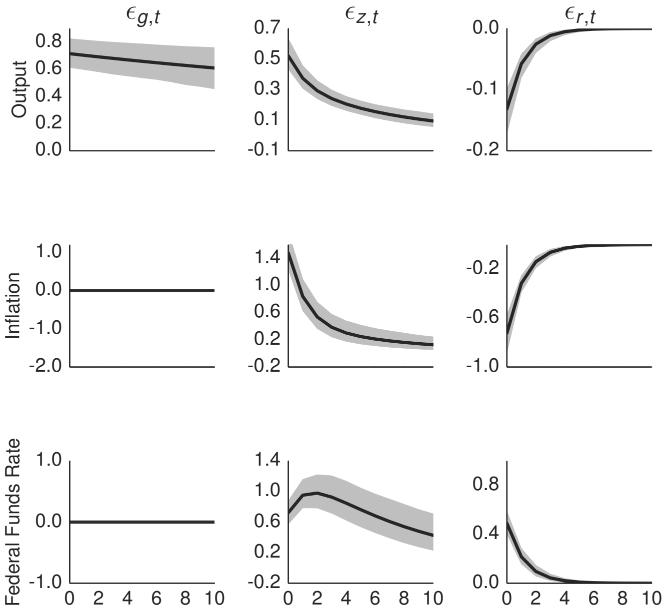
Example 1: New Keynesian DSGE

- We can solve the New Keynesian DSGE model described earlier.
- Obtain state space representation

Observables

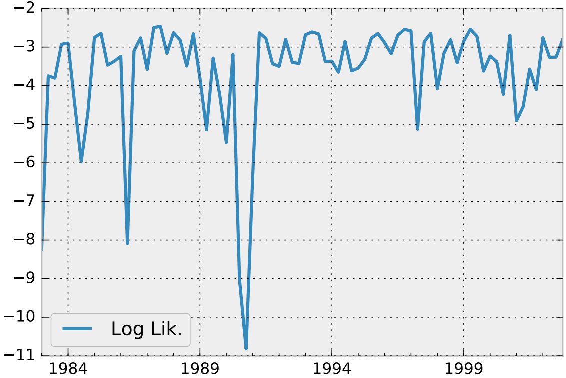


Impulse Responses

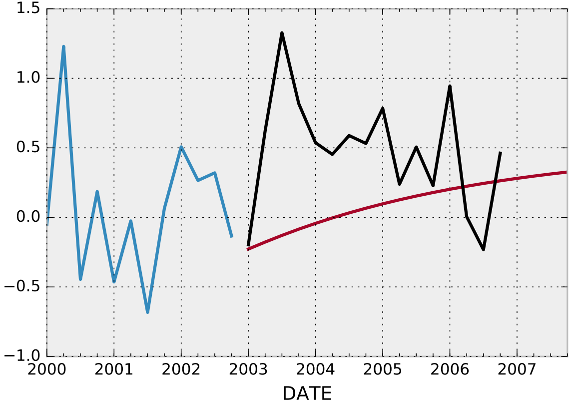


Filtered Technology Shock (Mean)

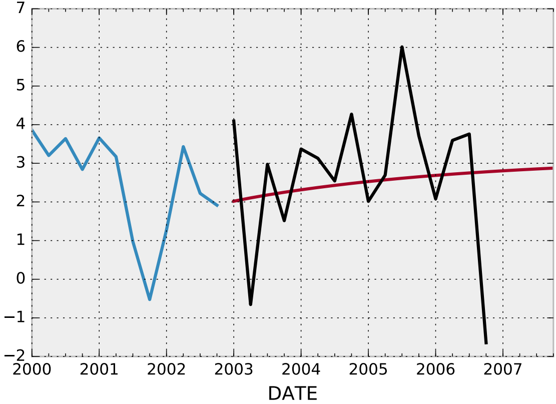
Log Likelihood Increments



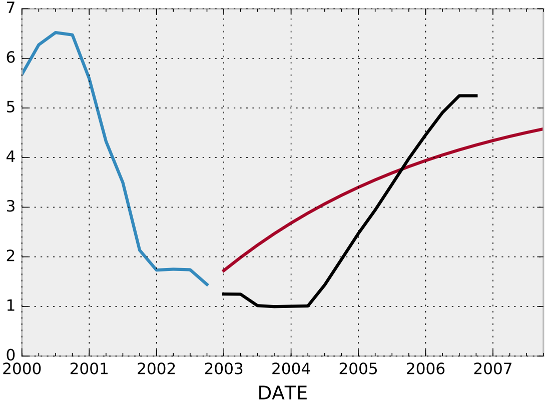
Forecast of Output Growth



Forecast of Inflation



Forecast of Interest Rate



References I

An, S. and Schorfheide, F. (2007). Bayesian analysis of dsge models. *Econometric Reviews*, 26(2-4):113–172.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Fernandez-Villaverde, J., Rubio-Ramirez, J., and Schorfheide, F. (2016). Solution and estimation methods for dsge models. *Handbook of Macroeconomics*, page 527 724.

Fernndez-Villaverde, J., Rubio-Ramrez, J. F., Sargent, T. J., and Watson, M. W. (2007). Abcs (and ds) of understanding vars. *American Economic Review*, 97(3):10211026.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press, 2nd edition.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

References II

Harvey, A. C. (1991). *Forecasting, Structural Time Series Models and the Kalman Filter*. University of Cambridge Press, Cambridge, United Kingdom.

Herbst, E. and Schorfheide, F. (2015). *Bayesian Estimation of DSGE Models*. Princeton University Press, Princeton.

Ireland, P. N. (2004). A method for taking models to the data. *Journal of Economic Dynamics and Control*, 28(6):1205–1226.

Sims, C. A. (2002). Solving linear rational expectations models. *Computational Economics*, 20:1–20.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97:586–608.

Woodford, M. (2003). *Interest and Prices*. Princeton University Press.