EC 831: Empirical Methods in Macroeconomics

Aeimit Lakdawala

Michigan State University

Linear State Space Models

Observation Equation

$$y_t = Ax_t + H\xi_t + w_t$$

y_t, w_t are n x 1,
ξ_t is r x 1 and
x_t is k x 1

State Equation

$$\xi_{t+1} = F\xi_t + v_{t+1}$$

$$\begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim iiN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\right)$$

Observed variables: y_t,x_t
Unobserved variables: ξ_t,v_t and w_t

Consider an AR(1)

$$\hat{y}_t = \rho \hat{y}_{t-1} + \varepsilon_t$$

Consider an AR(1)

$$\hat{y}_t = \rho \hat{y}_{t-1} + \varepsilon_t$$

Define $\xi_t = \hat{y}_t$, $v_t = \varepsilon_t$ and $y_t = \hat{y}_t$, $w_t = 0$, $x_t = 0$ and $F = \rho$, $H = 1$ and $A = 0$

Observation Equation:

$$\begin{aligned} y_t &= Ax_t + H\xi_t + w_t \\ \hat{y}_t &= \hat{y}_t \end{aligned}$$

State Equation

$$\begin{aligned} \xi_{t+1} &= F\xi_t + v_{t+1} \\ \hat{y}_{t+1} &= \rho \hat{y}_t + \varepsilon_{t+1} \end{aligned}$$

Consider a MA(1):

$$\hat{y}_t = \mu + u_t + \theta u_{t-1}$$

Define $\xi_t = (u_t, u_{t-1})$, $y_t = \hat{y}_t$, $x_t = 1$, $w_t = 0$,$v_t = [u_t, 0]'$

Observation Equation:

$$\begin{aligned} y_t &= Ax_t + H\xi_t + w_t \\ \hat{y}_t &= \mu + \begin{bmatrix} 1 & \theta \end{bmatrix} \begin{pmatrix} u_t \\ u_{t-1} \end{pmatrix} \end{aligned}$$

State Equation

$$\begin{aligned} \xi_{t+1} &= F\xi_t + v_{t+1} \\ \begin{pmatrix} u_{t+1} \\ u_t \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_t \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} u_{t+1} \\ 0 \end{pmatrix} \end{aligned}$$

Generally any ARMA(p,q) model can be put in state space form.

Solutions to linearized DSGE models can also be put in state space form.

Kalman Filter

The goal is to find the distribution of ξ_t given all information available upto time t, $\Omega_t = y_t, y_{t-1}, ..., y_1, x_t, x_{t-1}, ..., x_1$

$$p(\xi_t | \Omega_t)$$

Kalman Filter

The goal is to find the distribution of ξ_t given all information available upto time t, $\Omega_t = y_t, y_{t-1}, ..., y_1, x_t, x_{t-1}, ..., x_1$

$$p(\xi_t | \Omega_t)$$

Assume $\xi_0 \sim N(\xi_{0|0}, P_{0|0})$

- $\xi_{0|0}$: Best guess of ξ_0
- $P_{0|0}$: Uncertainty about this guess

Kalman Filter

The goal is to find the distribution of ξ_t given all information available upto time t, $\Omega_t = y_t, y_{t-1}, ..., y_1, x_t, x_{t-1}, ..., x_1$

$$p(\xi_t|\Omega_t)$$

Assume $\xi_0 \sim N(\xi_{0|0}, P_{0|0})$

- $\xi_{0|0}$: Best guess of ξ_0
- $P_{0|0}$: Uncertainty about this guess

With normal errors and linear structure we get

$$\xi_t|\Omega_t \sim N(\xi_{t|t}, P_{t|t})$$

Goal is to find $\xi_{t|t}$ and $P_{t|t}$

Kalman Filter

Two main steps:

1. Predict: Using all information upto time $t-1$ we want to obtain an optimal forecast of $\tilde{\zeta}_t$, let's call it $\tilde{\zeta}_{t|t-1}$.
2. Update: Once y_t is realized, we update our forecast, call it $\tilde{\zeta}_{t|t}$

Intuition for Kalman Filter

Assume z_1 and z_2 are jointly normal.

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \right)$$

Suppose now you observe z_1.
Now what is the conditional distribution of z_2?

$$z_2 | z_1 \sim N(m_{z_2}, V_{z_2})$$

Intuition for Kalman Filter

Assume z_1 and z_2 are jointly normal.

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \right)$$

Suppose now you observe z_1.
Now what is the conditional distribution of z_2?

$$z_2 | z_1 \sim N(m_{z_2}, V_{z_2})$$

$$m_{z_2} = \mu_2 + \Omega_{21}\Omega_{11}^{-1}\left[\mu_1 - z_1\right]$$

$$V_{z_2} = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$$

1st Predict step

$$
\begin{aligned}
\xi_1|\Omega_0 &\sim N(\xi_{1|0}, P_{1|0}) \\
\xi_{1|0} &= F\xi_{0|0} \\
P_{1|0} &= FP_{0|0}F' + Q
\end{aligned}
$$

1st Update step

$$\begin{bmatrix} y_1 | x_1, \Omega_0 \\ \xi_1 | x_1, \Omega_0 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \right)$$

where

$$\begin{aligned} \mu_1 &= Ax_1 + H\xi_{1|0} \\ \mu_2 &= \xi_{1|0} \\ \Omega_{11} &= HP_{1|0}H' + R \\ \Omega_{21} &= P_{1|0}H' \\ \Omega_{22} &= P_{1|0} \end{aligned}$$

1st Update step

$$\xi_1|y_1, x_1, \Omega_0 = \xi_1|\Omega_1 \sim N(\xi_{1|1}, P_{1|1})$$

$$\begin{aligned} \xi_{1|1} &= \xi_{1|0} + K\eta_{1|0} \\ P_{1|1} &= P_{1|0} - KHP_{1|0} \end{aligned}$$

where

$$\begin{aligned} K &= P_{1|0}H'(HP_{1|0}H' + R)^{-1} \\ \eta_{1|0} &= y_1 - Ax_1 - H\xi_{1|0} \end{aligned}$$

Kalman Filter Recursions

Predict:

$$\begin{aligned} \xi_{t|t-1} &= F\xi_{t-1|t-1} \\ P_{t|t-1} &= FP_{t-1|t-1}F' + Q \end{aligned}$$

Update:

$$\begin{aligned} \xi_{t|t} &= \xi_{t|t-1} + K\eta_{t|t-1} \\ P_{t|t} &= P_{t|t-1} - KHP_{t|t-1} \end{aligned}$$

where

$$\begin{aligned} \eta_{t|t-1} &= y_t - Ax_t - H\xi_{t|t-1} \\ f_{t|t-1} &= HP_{t|t-1}H' + R \\ K &= P_{t|t-1}H'f_{t|t-1}^{-1} \end{aligned}$$

How do we specify $\xi_{0|0}$ and $P_{0|0}$?

$$\xi_{t+1} = F\xi_t + v_t$$

If this is stationary, i.e. eigenvalues of F are inside the unit circle, then

$$\begin{aligned} \xi_{0|0} &= E(\xi_0) = 0 \\ P_{0|0} &= E(\xi_0\xi_0') \end{aligned}$$

$$vec(P_{0|0}) = (I - F \otimes F)^{-1} vec(Q)$$

If not, set diagonal elements of $P_{0|0}$ to large numbers

Bayesian Perspective

$\xi_0 \sim N(\xi_{0|0}, P_{0|0})$ is the prior.

$\xi_{t|t}$ is the posterior Bayesian expectation of $\xi_t|\Omega_t$ for given values of F,Q,A,H,R

So how do we estimate the parameters?

Parameter Estimation from classical perspective

What is the distribution of $y_t|\Omega_{t-1}, x_t, \theta$ where θ contains all the parameters.

$$y_t|\Omega_{t-1}, x_t, \theta \sim N(y_{t|t-1}, f_{t|t-1})$$

where

$$\begin{aligned} y_{t|t-1} &= Ax_t + H\xi_{t|t-1} \\ f_{t|t-1} &= HP_{t|t-1}H' + R \end{aligned}$$

Likelihood (based on "Prediction Error Decomposition")

$$lnL = -\frac{1}{2}\sum_{t=1}^{T} ln(2\pi f_{t|t-1}) - \frac{1}{2}\sum_{t=1}^{T} \eta'_{t|t-1} f^{-1}_{t|t-1} \eta_{t|t-1}$$

Parameter Estimation from classical perspective

$$lnL = -\frac{1}{2}\sum_{t=1}^{T} ln(2\pi f_{t|t-1}) - \frac{1}{2}\sum_{t=1}^{T} \eta'_{t|t-1} f^{-1}_{t|t-1} \eta_{t|t-1}$$

Note $f_{t|t-1}$ and $\eta_{t|t-1}$ are functions of θ

Maximize the log-likelihood w.r.t θ.

Note $\xi_{t|t}$ is our best guess given all data upto time t.

But as econometricians we observe $T \geq t$

- Can we improve our best guess using all available data?

Note $\xi_{t|t}$ is our best guess given all data upto time t.

But as econometricians we observe $T \geq t$

- Can we improve our best guess using all available data?

Yes!

We can use "smoothed" inference.

$$\xi_t | \Omega_T \sim N(\xi_{t|T}, P_{t|T})$$

Same idea used as Kalman filter derivations.

Smoothing

$$
\begin{aligned}
\xi_{t|T} &= \xi_{t|t} + J_t \left(\xi_{t+1|T} - \xi_{t+1|t} \right) \\
P_{t|T} &= P_{t|t} + J_t (P_{t+1|T} - P_{t+1|t}) J_t' \\
J_t &= P_{t|t} F' P_{t+1|t}^{-1}
\end{aligned}
$$

For smoothed inference.

1. Filter forward to get $\xi_{t|t}$ and $P_{t|t}$
2. Smooth backward using above formulae to get $\xi_{t|T}$ and $P_{t|T}$

What if $P_{t+1|t}$ is singular?

See Durbin & Koopman (2002)
Time Series Analysis by State Space Methods

- Alternative formulation of Kalman filter recursion in Ch 4.
- Does not require inversion of $P_{t+1|t}$ for smoothing
- Their method is also computationally more efficient