Kalman Smoother

Lawrence J. Christiano

Background

- These notes follow the notes on the Kalman filter.
- The state space-observer system is:

$$
\begin{aligned}
\xi_t &= F\xi_{t-1} + u_t, \; Eu_t u_t' = Q, \\
Y_t &= H\xi_t + w_t, \; Ew_t w_t' = R,
\end{aligned}
$$

  where

  - w_t, u_s are iid over time and uncorrelated for all s, t.
  - u_t's uncorrelated with past ξ_t's.
  - w_t's uncorrelated with ξ_t's at all leads and lags.
  - Y_t's demeaned observed data.
  - Eigenvalues of F all less than 1 in absolute value.

Background

- Definitions:

$$Y_{t+j|t'} \equiv P\left[Y_{t+j}|\mathcal{Y}_{t'}\right],\ \xi_{t+j|t'} \equiv P\left[\xi_{t+j}|\mathcal{Y}_{t'}\right],\ j > 0,$$

  where $\mathcal{Y}_t \equiv [Y_1, ..., Y_t]$ and $t, t' = 1, ..., T.$
- Also,

$$P_{t+1|t'} \equiv E\left[\xi_{t+1} - \xi_{t+1|t'}\right]\left[\xi_{t+1} - \xi_{t+1|t'}\right]'.$$

- We assume that the following objects are in hand:

$$\xi_{t|t}, \xi_{t+1|t}, P_{t|t}, P_{t+1|t},$$

  for $t = 1, 2, ..., T.$
- Here, we solve for

$$\xi_{t|T}, P_{t|T},$$

  for $t = 1, 2, ..., T.$

Backward Pass

- The computations are recursive, beginning at the end of the data set, at time T.
- Because they go 'backwards' in time, the calculations we are about to derive are referred to as the *backward pass* of the Kalman smoother.
  - Because the output of the Kalman filter is an essential input to the computations, and those calculations operate recursively beginning at the start of the data set, they are referred to as the *forward pass*.
- The focus of the analysis is on what will initially seem like a weird expression,

  $$P\left(\xi_t | \xi_{t+1}, \mathcal{Y}_t\right),$$

  i.e., the projection of the state, ξ_t, on the state in the next period and on observations, $\mathcal{Y}_t$, up to and including period t.
- Although the expression we focus on is an odd one, it will become clear that this is very convenient.

Using Recursive Property of Projections

- Recursive property of projections (derived in the notes, 'Properties of Projections'),

$$
\begin{aligned}
P\left(\xi_t | \xi_{t+1}, \mathcal{Y}_t\right) &= \xi_{t|t} + P\left[\xi_t - \xi_{t|t} | \xi_{t+1} - \xi_{t+1|t}\right] \\
&= \xi_{t|t} + J_t\left(\xi_{t+1} - \xi_{t+1|t}\right),
\end{aligned}
$$

  where J_t solves

$$
\min_{J_t} E\left[\xi_t - \xi_{t|t} - J_t\left(\xi_{t+1} - \xi_{t+1|t}\right)\right]^2
$$

First order condition:

$$
E\left[\xi_t - \xi_{t|t} - J_t\left(\xi_{t+1} - \xi_{t+1|t}\right)\right]\left(\xi_{t+1} - \xi_{t+1|t}\right)' = 0
$$

- Using

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

the first order condition

$$\begin{aligned} & E\left(\xi_t - \xi_{t|t}\right)\left(\xi_{t+1} - \xi_{t+1|t}\right)' \\ = \;& J_t E\left(\xi_{t+1} - \xi_{t+1|t}\right)\left(\xi_{t+1} - \xi_{t+1|t}\right)' \end{aligned}$$

reduces to:

$$\overbrace{E\left(\xi_t - \xi_{t|t}\right)\left(F\left(\xi_t - \xi_{t|t}\right) + u_{t+1}\right)'}^{=P_{t|t}F'}$$

$$= J_t \overbrace{E\left(\xi_{t+1} - \xi_{t+1|t}\right)\left(\xi_{t+1} - \xi_{t+1|t}\right)'}^{=P_{t+1|t}}$$

$$= 0.$$

or,

$$J_t = P_{t|t}F'P_{t+1|t}^{-1}.$$

Orthogonality

- We conclude that

$$P\left(\xi_t|\xi_{t+1},\mathcal{Y}_t\right)=\xi_{t|t}+\overbrace{P_{t|t}F'P_{t+1|t}^{-1}}^{J_t}\left(\xi_{t+1}-\xi_{t+1|t}\right),$$

- We now establish the following result:

$$P\left(\xi_t|\xi_{t+1},\mathcal{Y}_t\right)=P\left(\xi_t|\xi_{t+1},\mathcal{Y}_T\right).$$

  That is, there is no information in $\xi_{t+1},\mathcal{Y}_T$ that cannot be found in $\xi_{t+1},\mathcal{Y}_t$.

- We establish the result by using orthogonality. That is, we show that

$$\xi_t-P\left(\xi_t|\xi_{t+1},\mathcal{Y}_t\right)\perp y_{t+1},y_{t+2},...,y_T,$$

  so that the optimization problem that defines $P\left(\xi_t|\xi_{t+1},\mathcal{Y}_T\right)$ assigns zero weights to each of $y_{t+1},y_{t+2},...,y_T$.

Using Orthogonality

- We want to establish the following result:

$$\xi_t - P(\xi_t | \xi_{t+1}, \mathcal{Y}_t) \perp y_{t+1}, y_{t+2}, ..., y_T.$$

- Consider

$$y_{t+1} = H\xi_{t+1} + w_{t+1}.$$

Orthogonality of y_{t+1} follows from the fact that it is a linear combination of ξ_{t+1} and w_{t+1} and

$$\xi_t - P(\xi_t | \xi_{t+1}, \mathcal{Y}_t) \perp \xi_{t+1}, w_{t+1}.$$

Using Orthogonality

- Note that

$$
\begin{aligned}
y_{t+2} &= H\xi_{t+2} + w_{t+2} \\
&= HF\xi_{t+1} + Hu_{t+2} + w_{t+2} \\
&\quad \cdots \\
y_{t+j} &= HF^{j-1}\xi_{t+1} + Hu_{t+j} + HFu_{t+j-1} + ... + HF^{j-2}u_{t+2} + w
\end{aligned}
$$

  for $j = 2, 3, ...$.

- Orthogonality of y_{t+j}, $j \geq 2$ follows from the fact that y_{t+j} are linear functions of

$$
\xi_{t+1}, u_{t+2}, ..., u_{t+j}, w_{t+2}, ..., w_{t+j},
$$

  each of which is orthogonal to

$$
\xi_t - P\left(\xi_t | \xi_{t+1}, \mathcal{Y}_t\right).
$$

- We conclude that

$$
P\left(\xi_t | \xi_{t+1}, \mathcal{Y}_t\right) = P\left(\xi_t | \xi_{t+1}, \mathcal{Y}_T\right).
$$

Using Law of Iterated Projections

- Law of Iterated Projections (see 'Properties of Projections', LIP):

$$\begin{aligned} P\left[\xi_t|\mathcal{Y}_T\right] \overbrace{=}^{LIP} & P\left[P\left(\xi_t|\xi_{t+1},\mathcal{Y}_T\right)|\mathcal{Y}_T\right] \\ = & P\left[\xi_{t|t}+J_t\left(\xi_{t+1}-\xi_{t+1|t}\right)|\mathcal{Y}_T\right] \\ = & \xi_{t|t}+J_t\left(\xi_{t+1|T}-\xi_{t+1|t}\right). \end{aligned}$$

  We have used that

$$P\left[\xi_{t|t}|\mathcal{Y}_T\right]=\xi_{t|t},$$

  because $\xi_{t|t}$ can be constructed exactly from the data in $\mathcal{Y}_T$. The same is true for ξ_{t+1+T}.

- We are now in a position to compute what we're after, using

$$P\left[\xi_t|\mathcal{Y}_T\right]=\xi_{t|t}+J_t\left(\xi_{t+1|T}-\xi_{t+1|t}\right)$$

  for $t=1,2,...,T$.

Bottom Line

- The following equation can be solved recursive, 'backwards'.
- For $t = T$, we obtain

$$\xi_{T|T} = P\left[\xi_T | \mathcal{Y}_T\right]$$

  from the forward pass using the Kalman filter.
- For $t = T - 1$,

$$\xi_{T-1|T} = \xi_{T-1|T-1} + J_{T-1}\left(\xi_{T|T} - \xi_{T|T-1}\right),$$

  can be solved for $\xi_{T-1|T}$ given the objects, $\xi_{T-1|T-1}$, J_{T-1} and $\xi_{T|T-1}$ found on the forward pass.
- Similarly, for $t = T - 2, T - 3, ..., 1$:

$$\begin{aligned}
\xi_{T-2|T} &= \xi_{T-2|T-2} + J_{T-2}\left(\xi_{T-1|T} - \xi_{T-1|T-2}\right) \\
\xi_{T-3|T} &= \xi_{T-3|T-3} + J_{T-3}\left(\xi_{T-2|T} - \xi_{T-2|T-3}\right) \\
& \dots \\
\xi_{1|T} &= \xi_{1|1} + J_t\left(\xi_{2|T} - \xi_{2|1}\right)
\end{aligned}$$