Kalman Filter

Lawrence J. Christiano

Background

- The Kalman filter is a powerful tool, which can be used in a variety of contexts.
  - can be used for filtering and smoothing.
- To help make it concrete, we will derive the filter here.
  - basic tool for forecasting, and for computing forecast confidence intervals.

State Space/Observer Form

- Canonical representation of data:

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

  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.
  - Eigenvalues of F all less than 1 in absolute value.

- Let Y_t denote demeaned data, $Y_t \equiv Y_t^{data} - a$,
- Will derive the *Kalman filter*, which solves the *projection problem*:

$$
Y_{t+j|t} \equiv P\left[Y_{t+j}|\mathcal{Y}_t\right],\ \xi_{t+j|t} \equiv P\left[\xi_{t+j}|\mathcal{Y}_t\right],\ j > 0
$$

  where $\mathcal{Y}_t \equiv [Y_1, ..., Y_t]$. We simplify by setting $w_t = 0$ for all t.

Example of Projection

- Let the log wage rate, w, and log price level, p, be

$$\begin{aligned} w &= z+u \\ p &= z+v, \end{aligned}$$

  where u and v are uncorrelated with each other and with z. All have zero mean.

- Suppose you observe w, but what you're really interested in is $w-p$.
  - obviously a move in w that reflects z is not interesting to you.

- You form the projection,

$$P\left[w-p|w\right] \equiv \alpha w,$$

  where α solves

$$\min_{\alpha} E\left[w-p-\alpha w\right]^2$$

Orthogonality Property of Projections:

- Projection solves a particular optimization problem:

$$\min_{\alpha} E\left[w - p - \alpha w\right]^2$$

- First order condition:

$$\begin{aligned} E\overbrace{[w - p - \alpha w]}^{\text{projection error}}w &= 0 \rightarrow \\ \alpha &= \frac{E\left(w-p\right)w}{Ew^2} \\ &= \frac{E\left(u-v\right)\left(z+u\right)}{E\left(z+u\right)^2} \\ &= \frac{\sigma_u^2}{\sigma_u^2+\sigma_z^2} = \frac{\sigma_u^2/\sigma_z^2}{\sigma_u^2/\sigma_z^2+1} \end{aligned}$$

- Orthogonality of projections: projection error uncorrelated with information, w, used in computing the projection.

The Filter

- Will compute projections:

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

recursively:

$$\left(\xi_{1|0}, Y_{1|0}\right), \left(\xi_{2|1}, Y_{2|1}\right), ..., \left(\xi_{T+1|T}, Y_{T+1|T}\right)$$

- Will simultaneously compute measures of uncertainty:

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

Forecasts of the Data

- Will focus primarily on forecasting ξ_t because forecasts of Y_t easy to read from forecast of ξ_t

$$
\begin{aligned}
P\left[Y_{t+j}|\mathcal{Y}_t\right] &= P\left[H\xi_{t+j} + \overbrace{w_{t+j}}^{\text{uncorrelated with everything in } \mathcal{Y}_t} |\mathcal{Y}_t\right] \\
&= HP\left[\xi_{t+j}|\mathcal{Y}_t\right] + P\left[w_{t+j}|\mathcal{Y}_t\right] = H\xi_{t+j|t}.
\end{aligned}
$$

- Also,

$$
\begin{aligned}
& E\left[Y_{t+j} - Y_{t+j|t}\right]\left[Y_{t+j} - Y_{t+j|t}\right]' \\
= \;& E\left[H\xi_{t+j} + w_{t+j} - H\xi_{t+j|t}\right]\left[H\xi_{t+j} + w_{t+j} - H\xi_{t+j|t}\right]' \\
= \;& E\left[H\left(\xi_{t+j} - \xi_{t+j|t}\right) + w_{t+j}\right]\left[\left(\xi_{t+j} - \xi_{t+j|t}\right)H' + w_{t+j}\right]' \\
= \;& HP_{t+j||t}H' + R.
\end{aligned}
$$

First Date of the Filter

- At $t = 0$ have $\mathcal{Y}_0 = \phi$, the empty set.
- So,

$$\xi_{1|0} = P[\xi_1|\mathcal{Y}_0] = E\xi_1 = 0,$$

  the unconditional expectation. Also,

$$P_{1|0} = E\left[\xi_1 - \xi_{1|0}\right]\left[\xi_1 - \xi_{1|0}\right]' = V,$$

  say, where V denotes the variance of ξ_1.
- Compute V by solving the Ricatti equation:

$$V = E[F\xi_{t-1} + u_t][F\xi_{t-1} + u_t]' = FVF' + Q$$

- Most robust way to find V is $V = V_\infty$ in:
  - Set V_0 to be *any* pos. def. matrix, compute $V_{j+1} = FV_jF' + Q,\ j = 0, 1, 2, \ldots$

An Intermediate Date with the Filter

- Suppose we have $\xi_{t|t-1}, P_{t|t-1}$ in hand.

- We now receive a new observation, Y_t.

- Want to compute

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

- We do this in two steps:
  - First, compute $\xi_{t|t}, P_{t|t}$. Second, compute $\xi_{t+1|t}, P_{t+1|t}$.

First Step for the Filter

- Basic recursive property of projections:

$$\xi_{t|t} = \xi_{t|t-1} + P\left[\overbrace{\xi_t - \xi_{t|t-1}}^{\text{forecast error in } \xi_{t|t-1}} \quad | \quad \overbrace{Y_t - \underbrace{H\xi_{t|t-1}}_{\equiv Y_{t|t-1}}}^{\text{new information in } Y_t \text{ not in } \mathcal{Y}_{t-1}}\right]$$

- This formula is obviously ‘correct’ in the special case where the information in Y_t allows you to compute the forecast error, $\xi_t - \xi_{t|t-1}$, exactly.
- Has a learning interpretation
  - you update your old guess, $\xi_{t|t-1}$, about ξ_t using what is new about the information in Y_t, i.e., using $Y_t - H\xi_{t|t-1}$.

First Step for the Filter

- Write

$$\begin{aligned} \xi_{t|t} &= \xi_{t|t-1} + P\left[\xi_t - \xi_{t|t-1} | Y_t - H\xi_{t|t-1}\right] \\ &= \xi_{t|t-1} + \alpha_t \left[Y_t - H\xi_{t|t-1}\right], \end{aligned}$$

where the matrix, α_t, solves

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

- First order condition associated with optimality:

$$E\left[\xi_t - \xi_{t|t-1} - \alpha_t \left(Y_t - H\xi_{t|t-1}\right)\right]\left[Y_t - H\xi_{t|t-1}\right]' = 0,$$

which again is the orthogonality of projections.

First Step for the Filter

- First order condition implies:

$$\begin{aligned} & E\left[\xi_t - \xi_{t|t-1}\right]\left[Y_t - H\xi_{t|t-1}\right]' \\ = \; & \alpha_t E\left(Y_t - H\xi_{t|t-1}\right)\left(Y_t - H\xi_{t|t-1}\right)' \end{aligned}$$

or,

$$\begin{aligned} & \overbrace{E\left[\xi_t - \xi_{t|t-1}\right]\left[\xi_t - \xi_{t|t-1}\right]'}^{P_{t|t-1}} H' \\ = \; & \alpha_t H \overbrace{E\left(\xi_t - \xi_{t|t-1}\right)\left(\xi_t - \xi_{t|t-1}\right)'}^{P_{t|t-1}} H', \end{aligned}$$

so that

$$\alpha_t = P_{t|t-1} H' \left(H P_{t|t-1} H'\right)^{-1}.$$

First Step for the Filter

- We conclude

$$\xi_{t|t} = \xi_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H'\right)^{-1}\left[Y_t - H\xi_{t|t-1}\right].$$

- With $\xi_{t|t}$ in hand, we move on to $P_{t|t}$:

$$\begin{aligned}
P_{t|t} &= E\left[\xi_t - \xi_{t|t}\right]\left[\xi_t - \xi_{t|t}\right]' \\
&= E\left[\overbrace{\xi_t - \xi_{t|t-1} - \alpha_t\left(Y_t - H\xi_{t|t-1}\right)}^{\text{orthogonal to }\left(Y_t - H\xi_{t|t-1}\right)}\right] \\
&\quad \times \left[\xi_t - \xi_{t|t-1} - \alpha_t\left(Y_t - H\xi_{t|t-1}\right)\right]' \\
&= E\left[\xi_t - \xi_{t|t-1} - \alpha_t\left(Y_t - H\xi_{t|t-1}\right)\right]\left[\xi_t - \xi_{t|t-1}\right]',
\end{aligned}$$

by orthogonality.

First Step for the Filter

- From the previous slide,

$$
\begin{aligned}
P_{t|t} &= E\left[\xi_t - \xi_{t|t-1} - \alpha_t\left(Y_t - H\xi_{t|t-1}\right)\right]\left[\xi_t - \xi_{t|t-1}\right]' \\
&= P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H'\right)^{-1}HP_{t|t-1}
\end{aligned}
$$

  completing the derivation of $\xi_{t|t}$ and $P_{t|t}$.

- Now we proceed to the second step, to compute $\xi_{t+1|t}$ and $P_{t+1|t}$.

Second Step for the Filter

- By linearity of projections:

$$\xi_{t+1|t} = F\xi_{t|t} + \overbrace{u_{t+1|t}}^{=0}.$$

- It follows that:

$$\begin{aligned}\xi_{t+1|t} &= \overbrace{F\xi_{t|t-1}}^{\text{forecast, } \xi_{t+1|t-1}\text{, based on } t-1 \text{ info, } \mathcal{Y}_{t-1}} \\ &+ \overbrace{FP_{t|t-1}H'\left(HP_{t|t-1}H'\right)^{-1}}^{\text{Kalman gain matrix, } K_t}\overbrace{\left[Y_t - H\xi_{t|t-1}\right]}^{\text{new information}}.\end{aligned}$$

- Next, $P_{t+1|t}$....

Second Step for the Filter

- Finally,

$$\begin{aligned} P_{t+1|t} &= E\left[\xi_{t+1} - \xi_{t+1|t}\right]\left[\xi_{t+1} - \xi_{t+1|t}\right]' \\ &= E\left[F\left(\xi_t - \xi_{t|t}\right) + u_{t+1}\right]\left[F\left(\xi_t - \xi_{t|t}\right) + u_{t+1}\right]' \\ &= FP_{t|t}F' + Q \\ &= F\left[P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H'\right)^{-1}HP_{t|t-1}\right]F' + Q. \end{aligned}$$

- Done! We now have

$$\left(\xi_{1|0}, P_{1|0}\right), ..., \left(\xi_{T+1|T}, P_{T+1|T}\right)$$

and also

$$\left(\xi_{1|1}, P_{1|1}\right), ..., \left(\xi_{T|T}, P_{T|T}\right)$$

Forecasting

- We have the one-step-ahead forecast and its uncertainty:

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

- Then,

$$\xi_{T+2|T} = P\left[\xi_{T+2}|\mathcal{Y}_t\right] = \overbrace{FP\left[\xi_{T+1}|\mathcal{Y}_T\right]}^{=F\xi_{T+1|T}} + \overbrace{P\left[u_{T+2}|\mathcal{Y}_T\right]}^{=0}$$

and so on:

$$\xi_{T+j|T} = F^{j-1}\xi_{T+1|T}.$$

Forecasting

- Want measures of forecast uncertainty.
- For $T+2$:

$$\begin{aligned} P_{T+2|T} &= E\left[\xi_{T+2}-\xi_{T+2|T}\right]\left[\xi_{T+2}-\xi_{T+2|T}\right]' \\ &= E\left[F\left(\xi_{T+1}-\xi_{T+1|T}\right)+u_{T+2}\right]\left[F\left(\xi_{T+1}-\xi_{T+1|T}\right)\right. \\ &= FP_{T+1|T}F'+Q \end{aligned}$$

- Similarly, for $j>1$

$$\begin{aligned} P_{T+j|T} &= E\left[\xi_{T+j}-\xi_{T+j|T}\right]\left[\xi_{T+j}-\xi_{T+j|T}\right]' \\ &= E\left[F\left(\xi_{T+j}-\xi_{T+j|T}\right)+u_{T+j}\right] \\ &\quad \times\left[F\left(\xi_{T+j}-\xi_{T+j|T}\right)+u_{T+j}\right]' \\ &= FP_{T+j-1|T}F'+Q \end{aligned}$$

Forecasting

- Note, as $j \to \infty$,
  - $P_{T+j|T} \to V$
  - $\xi_{T+j|T} \to 0$
- These features follow from the fact that the eigenvalues of F are less than unity in absolute value.
- Message: for observations far in the future, available data not helpful and might as well just guess the uconditional mean, with forecast error variance equal to unconditionarly

Smoothing

- We have reviewed *filtering*, which is what is used in forecasting (and, calculation of likelihood).
- Also useful to do *smoothing*:

$$P\left[\xi_t|\mathcal{Y}_T\right],\ t=1,2,...,T.$$

  Smoothing gives the best guess about the value taken on by a variable that is in the model (like the output gap, or the natural rate of interest), but that is not contained among the observed data.

  - Derivations of the Kalman smoother first derive the Kalman filter, as we did, and then derive the smoother as a second step.