Kalman Filter

Juan F. Rubio-Ramírez

Emory University

State Space Form

- Transition equation:

$$x_{t+1} = Fx_t + G\omega_{t+1},\ \omega_{t+1} \sim N(0, Q)$$

- Measurement equation

$$z_t = H'x_t + \upsilon_t,\ \upsilon_t \sim N(0, R)$$

- Where x_t are the states and z_t are the observables
- Assume we want to write the likelihood function of $z^T = \{z_t\}_{t=1}^T$.
- A necessary and sufficient condition for the system to be stable is that:

$$|\lambda_i(F)| < 1$$

for all i, where $\lambda_i(F)$ stands for eigenvalue of F.

The State Space Representation is Not Unique

- Let the following system
  - Transition equation

$$x_{t+1} = Fx_t + G\omega_{t+1},\ \omega_{t+1} \sim N(0, Q)$$

  - Measurement equation

$$z_t = H'x_t + \upsilon_t,\ \upsilon_t \sim N(0, R)$$

- Let B be a non-singular squared matrix conforming with F. Then, if $x_t^* = Bx_t$, $F^* = BFB^{-1}$, $G^* = BG$, and $H^* = (H'B)'$, we can write:
  - Transition equation

$$x_{t+1}^* = F^* x_t^* + G^* \omega_{t+1},\ \omega_{t+1} \sim N(0, Q)$$

  - Measurement equation

$$z_t = H^{*\prime} x_t^* + \upsilon_t,\ \upsilon_t \sim N(0, R)$$

Example

- Assume the following AR(2) process $z_t = \rho_1 z_{t-1} + \rho_2 z_{t-2} + \upsilon_t,\ \upsilon_t \sim N\left(0, \sigma_\upsilon^2\right)$.
- State Space I
  - Transition equation

$$x_t = \begin{bmatrix} \rho_1 & 1 \\ \rho_2 & 0 \end{bmatrix} x_{t-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \upsilon_t$$

  - Where $x_t = \begin{bmatrix} y_t & \rho_2 y_{t-1} \end{bmatrix}'$ and measurement equation $z_t = \begin{bmatrix} 1 & 0 \end{bmatrix} x_t$
- State Space II
  - Transition equation

$$x_t = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} x_{t-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \upsilon_t$$

  - Where $x_t = \begin{bmatrix} y_t & y_{t-1} \end{bmatrix}'$ and measurement equation $z_t = \begin{bmatrix} 1 & 0 \end{bmatrix} x_t$
- Try $B = \text{diag}\begin{pmatrix} 1 & \rho_2 \end{pmatrix}$ on the second system to get the first system.

What is the Kalman Filter trying to do?

- Let $x_{t|t-1} = E\left(x_t | z^{t-1}\right)$.
- Let $z_{t|t-1} = E\left(z_t | z^{t-1}\right) = H' x_{t|t-1}$.
- Let $x_{t|t} = E\left(x_t | z^{t}\right)$.
- Let assume we have $x_{t|t-1}$ and $z_{t|t-1}$.
- We observe a new z_t.
- We need to obtain $x_{t|t}$.
- Note that $x_{t+1|t} = F x_{t|t}$ and $z_{t+1|t} = H' x_{t+1|t}$, so we can wait for z_{t+1}.
- Therefore, the key question is how to obtain from $x_{t|t-1}$ and z_t.

Kalman Filter Gain to find $x_{t|t}$

- Assume we use the following equation to get $x_{t|t}$ from z_t and $x_{t|t-1}$:

$$x_{t|t} = x_{t|t-1} + K_t \left(z_t - z_{t|t-1}\right) = x_{t|t-1} + K_t \left(z_t - H' x_{t|t-1}\right)$$

- This formula will have some probabilistic justification (to follow).
- K_t is called the Kalman filter gain and it measures how much we update $x_{t|t-1}$ as a function in our error in predicting z_t.
- Let $\Sigma_{t|t-1} = E\left(\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' |z^{t-1}\right)$.
- Let $\Omega_{t|t-1} = E\left(\left(z_t - z_{t|t-1}\right)\left(z_t - z_{t|t-1}\right)' |z^{t-1}\right)$.
- Let $\Sigma_{t|t} = E\left(\left(x_t - x_{t|t}\right)\left(x_t - x_{t|t}\right)' |z^{t}\right)$.
- Let $\Phi_{t|t-1} = E\left(\left(z_t - z_{t|t}\right)\left(x_t - x_{t|t}\right)' |z^{t-1}\right)$.
- It can be shown that:

$$K_t = \Sigma_{t|t-1} H \left(H' \Sigma_{t|t-1} H + R\right)^{-1}$$

Intuition

- The gain:

$$K_t = \Sigma_{t|t-1} H \left(H' \Sigma_{t|t-1} H + R \right)^{-1}$$

- Update of $x_{t|t}$ given z_t and $x_{t|t-1}$ being:

$$x_{t|t} = x_{t|t-1} + K_t \left(z_t - H' x_{t|t-1} \right)$$

- If we did a big mistake forecasting $x_{t|t-1}$ using past information ($\Sigma_{t|t-1}$ large) we give a lot of weight to the new information (K_t large).
- If the new information is noise (R large) we give a lot of weight to the old prediction (K_t small)

Example

- Assume the following model in State Space form:
  - Transition equation
  $$x_t = \mu + \upsilon_t,\ \upsilon_t \sim N\left(0, \sigma_\upsilon^2\right)$$
  - Measurement equation
  $$z_t = x_t + \xi_t,\ \xi_t \sim N\left(0, \sigma_\xi^2\right)$$
  - Let $\sigma_\xi^2 = q\sigma_\upsilon^2$.
- Then, if $\Sigma_{t|t-1} = \sigma_\upsilon^2$, we have
$$K_t = \sigma_\upsilon^2 \frac{1}{1+q} \propto \frac{1}{1+q}.$$
- The bigger σ_ξ^2 relative to σ_υ^2 (the bigger q) the lower K_t and the less we trust z_t.

The Algorithm

- Given $\Sigma_{t|t-1}$, z_t, and $x_{t|t-1}$, we can now set the Kalman filter algorithm:
- Let $\Sigma_{t|t-1}$, then we compute: $\Omega_{t|t-1} = E\left(\left(z_t - z_{t|t-1}\right)\left(z_t - z_{t|t-1}\right)' | z^{t-1}\right)$

$$= E\left(\begin{array}{c} H'\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' H + \\ \upsilon_t\left(x_t - x_{t|t-1}\right)' H + H'\left(x_t - x_{t|t-1}\right)\upsilon_t' + \\ \upsilon_t \upsilon_t' | z^{t-1} \end{array}\right) = H'\Sigma_{t|t-1}H + R$$

- Let $\Sigma_{t|t-1}$, then we compute: $\Phi_{t|t-1} = E\left(\left(z_t - z_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' | z^{t-1}\right)$

$$= E\left(H'\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' + \upsilon_t\left(x_t - x_{t|t-1}\right)' | z^{t-1}\right) = H'\Sigma_{t|t-1}$$

- Let $\Sigma_{t|t-1}$, then we compute $K_t = \Sigma_{t|t-1}H\left(H'\Sigma_{t|t-1}H + R\right)^{-1}$.
- Let $\Sigma_{t|t-1}$, $x_{t|t-1}$, K_t, and z_t then we compute $x_{t|t} = x_{t|t-1} + K_t\left(z_t - H'x_{t|t-1}\right)$

The Algorithm

- Let $\Sigma_{t|t-1}$ and K_t, then we compute: $\Sigma_{t|t} = E\left(\left(x_t - x_{t|t}\right)\left(x_t - x_{t|t}\right)' | z^t\right)$

$$= E\begin{pmatrix} \left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' - \\ \left(x_t - x_{t|t-1}\right)\left(z_t - H'x_{t|t-1}\right)' K_t' - \\ K_t\left(z_t - H'x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)' + \\ K_t\left(z_t - H'x_{t|t-1}\right)\left(z_t - H'x_{t|t-1}\right)' K_t' | z^t \end{pmatrix} = \Sigma_{t|t-1} - K_t H' \Sigma_{t|t-1}$$

- Let $\Sigma_{t|t}$, then we compute $\Sigma_{t+1|t} = F\Sigma_{t|t}F' + GQG'$
- Let $x_{t|t}$, then we compute:
  - $x_{t+1|t} = Fx_{t|t}$
  - $z_{t+1|t} = H'x_{t+1|t}$
- Therefore, from $x_{t|t-1}$, $\Sigma_{t|t-1}$, and z_t we compute $x_{t|t}$ and $\Sigma_{t|t}$.
- We also compute $z_{t|t-1}$, $\Omega_{t|t-1}$, and $\Phi_{t|t-1}$. Why for? As you will see later, they are necessary to calculate the likelihood function of $z^T = \{z_t\}_{t=1}^T$ (to follow).

Summary

We start with $x_{t|t-1}$ and $\Sigma_{t|t-1}$, the we observe z_t and....

- $\Omega_{t|t-1} = H'\Sigma_{t|t-1}H + R$
- $z_{t|t-1} = H'x_{t|t-1}$
- $K_t = \Sigma_{t|t-1}H\left(H'\Sigma_{t|t-1}H + R\right)^{-1}$
- $\Sigma_{t|t} = \Sigma_{t|t-1} - K_t H'\Sigma_{t|t-1}$
- $x_{t|t} = x_{t|t-1} + K_t\left(z_t - H'x_{t|t-1}\right)$
- $\Sigma_{t+1|t} = F\Sigma_{t|t}F' + GQG'$
- $x_{t+1|t} = Fx_{t|t}$

We finish with $x_{t+1|t}$ and $\Sigma_{t+1|t}$.

A Probabilistic Approach to K_t

- Assume
$$Z|w = [X'|w \;\; Y'|w]' \sim N\left(\begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}\right)$$
- then
$$X|y,w \sim N\left(\overline{x} + \Sigma_{xy}\Sigma_{yy}^{-1}\left(y - \overline{y}\right), \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}\right)$$
- But $x_{t|t-1} = E\left(x_t|z^{t-1}\right)$ and $\Sigma_{t|t-1} = E\left(\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)'|z^{t-1}\right)$
- If $z_t|z^{t-1}$ is the random variable z_t (observable) conditional on z^{t-1}, then:
  - Let $z_{t|t-1} = H'x_{t|t-1}$
  - Let $\Omega_{t|t-1} = H'\Sigma_{t|t-1}H + R$
  - Let $\Phi_{t|t-1} = H'\Sigma_{t|t-1}$

A Probabilistic Approach to K_t

- Assume we know $x_{t|t-1}$ and $\Sigma_{t|t-1}$, then

$$\left(\begin{array}{c} x_t \\ z_t \end{array} |z^{t-1} \right) N \left(\left[\begin{array}{c} x_{t|t-1} \\ H'x_{t|t-1} \end{array} \right], \left[\begin{array}{cc} \Sigma_{t|t-1} & \Sigma_{t|t-1}H \\ H'\Sigma_{t|t-1} & H'\Sigma_{t|t-1}H + R \end{array} \right] \right)$$

- Remember that:

$$X|y, w \sim N\left(\overline{x} + \Sigma_{xy}\Sigma_{yy}^{-1}\left(y - \overline{y}\right), \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}\right)$$

- Then, we can write $x_t|z_t, z^{t-1} = x_t|z^t \sim N\left(x_{t|t}, \Sigma_{t|t}\right)$, where

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

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \underbrace{\Sigma_{t|t-1}H\left(H'\Sigma_{t|t-1}H + R\right)^{-1}}_{K_t}H'\Sigma_{t|t-1}$$