Kalman Filter

Setup

Observable equation is described by

$$y_t = x_t + \eta_t \tag{1}$$

where y_t can be observed at time t and $\eta_t \sim N(0, \sigma_\eta^2)$ is a noise shock which prevent to correctly observe state x_t.

State transition equation is described by

$$x_t = x_{t-1} + \varepsilon_t \tag{2}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is a structural shock which affects the transition from x_{t-1} to x_t.

Procedure

Goal is to optimize the forecast of x_{t+1} using available information at time t, i.e. y_t. Assume an initial value for $x_{1,0}$ and $Var(x_{1,0} - x_{0,0}) = P_{1,0}$. Procedure can be summarized as follows,

1. Forecast y_t using information at time $t-1$ and evaluate the error variance of this prediction.
2. Infer x_t using information at time t and evaluate the error variance of this inference.
3. Forecast x_{t+1} using information at time t and evaluate the error variance of this prediction.
4. Find the steady state variance $P_{t,t}$.

Given $x_{1,0}$ the three initial steps are

1. $y_{1,0} = x_{1,0}$ and the forecast error variance is $\Omega^y_{1,0} = \sigma_\eta^2$.
2. $x_{1,1} = y_1$ and the forecast error variance is $P_{1,1} = \sigma_\eta^2$.
3. $x_{2,1} = x_{1,1}$ and the forecast error variance is $P_{2,1} = \sigma_\eta^2 + \sigma_\varepsilon^2$.

As a generalization, the three steps are

1. $y_{t,t-1} = x_{t,t-1}$ and the forecast error variance is $\Omega^y_{t,t-1} = Var(x_{t,t-1}) + \sigma_\eta^2 = P_{t,t-1} + \sigma_\eta^2$.
2. We want to forecast $x_{t,t}$ using all the available information up to time t. As a simplification we try to forecast $x_t - x_{t,t-1}$ using $y_t - x_{t,t-1}$. Coefficient β^{PROJ} is derived as follows

$$\begin{aligned}
\beta^{KG} &= \frac{Cov(x_t - x_{t,t-1}, y_t - x_{t,t-1})}{Var(y_t - x_{t,t-1})} \\
&= \frac{Cov(x_t - x_{t,t-1}, x_t + \eta_t - x_{t,t-1})}{Var(x_t + \eta_t - x_{t,t-1})} \\
&= \frac{P_{t,t-1}}{A^2(P_{t,t-1} + \sigma_\eta^2)} \\
&= \frac{P_{t,t-1}}{P_{t,t-1} + \sigma_\eta^2}
\end{aligned} \tag{3}$$

This implies that

$$x_{t,t} - x_{t,t-1} = \beta^{KG}(y_t - x_{t,t-1}) \tag{4}$$

which is

$$\begin{aligned} x_{t,t} &= x_{t,t-1} + \beta^{KG}(y_t - x_{t,t-1}) \\ &= (1-\beta^{KG})x_{t,t-1} + \beta^{KG} y_t \end{aligned} \tag{5}$$

and this is the reason why β^{KG} is called Kalman gain, i.e. it defines how much to weight new information at time t to infer $x_{t,t}$.

Now we need to figure out the forecast error variance of $x_t - x_{t,t}$, i.e. $P_{t,t} = Var(x_t - x_{t,t})$

$$\begin{aligned} P_{t,t} &= Var\Big[x_{t,t} - x_{t,t-1} - \beta^{KG}(y_t - x_{t,t-1})\Big] \\ &= Var\Big[x_{t,t} - x_{t,t-1} - \beta^{KG}(x_t + \eta_t - x_{t,t-1})\Big] \\ &= Var\Big[x_{t,t} - x_{t,t-1} - \frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2}(x_t + \eta_t - x_{t,t-1})\Big] \\ &= E\Big\{\Big[x_{t,t} - x_{t,t-1} - \frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2}(x_t + \eta_t - x_{t,t-1})\Big]^2\Big\} \\ &= E\Big\{\Big[x_{t,t} - x_{t,t-1}\Big]^2\Big\} - 2E\Big\{\Big(x_{t,t} - x_{t,t-1}\Big)\Big(\frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2}(x_t + \eta_t - x_{t,t-1})\Big)\Big\} \\ &+ E\Big\{\Big[\frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2}(x_t + \eta_t - x_{t,t-1})\Big]^2\Big\} \\ &= P_{t,t-1} - 2P_{t,t-1}\frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2} + P_{t,t-1}\frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2} \\ &= P_{t,t-1} - P_{t,t-1}\frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2} \\ &= P_{t,t-1}\Big[1 - \frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2}\Big] \\ &= P_{t,t-1}(1-\beta^{KG}) \end{aligned} \tag{6}$$

3. $x_{t+1,t} = x_{t,t}$ and the forecast error variance is $P_{t+1,t} = P_{t,t} + \sigma_\varepsilon^2$ which is

$$\begin{aligned} P_{t+1,t} &= P_{t,t} + \sigma_\varepsilon^2 \\ &= P_{t,t-1}\Big[1 - \frac{P_{t,t-1}}{P_{t,t-1}+\sigma_\eta^2}\Big] + \sigma_\varepsilon^2 \end{aligned} \tag{7}$$

4. Find the steady state value of P as follows

$$\begin{aligned} P &= P\Big[1 - \frac{P}{P+\sigma_\eta^2}\Big] + \sigma_\varepsilon^2 \\ &= P - \frac{P^2}{P+\sigma_\eta^2} + \sigma_\varepsilon^2 \end{aligned} \tag{8}$$

which is

$$P^2 + \sigma_\eta^2 P = P^2 + \sigma_\eta^2 P - P^2 + \sigma_\varepsilon^2 P + \sigma_\varepsilon^2 \sigma_\eta^2 \tag{9}$$

which is

$$P^2 - \sigma_\eta^2 P - \sigma_\varepsilon^2 \sigma_\eta^2 = 0 \tag{10}$$

which is

$$P = \frac{1}{2}\left(\sigma_\varepsilon^2 + \sqrt{\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 \sigma_\eta^2}\right) \tag{11}$$