The Kalman Filter

The Kalman filter is an Algorithm for sequentially updating a linear projection for a state-space representation.

Consider y_t a vector of of n variables observed at date t.

Then the state space representation can be written as

$$\begin{aligned} \xi_{t+1} &= F\xi_t + \upsilon_{t+1} \qquad \textit{State Equation} \\ y_t &= A'x_t + H'\xi_t + \omega_t \qquad \textit{Observation Equation} \end{aligned}$$

where F_{rxr}, A'_{nxk}, H'_{nxr} are matrices of parameters and $x_{t_{kx1}}$ is a vector of predetermined variables.

The shock υ_{t+1} is a white noise with

$$E(\upsilon_t \upsilon'_\tau) = \begin{cases} Q_{rxr} \text{ if } t = \tau \\ 0 \text{ Otherwise} \end{cases}$$

and ω_t is a white noise with

$$E(\omega_t \omega'_\tau) = \begin{cases} R_{nxn} \text{ if } t = \tau \\ 0 \text{ Otherwise} \end{cases} .$$

Since x_t is predetermined or exogenous, it does not provide information about ξ_{t+s} or ω_{t+s}, for s>0, beyond that contained in $\{y_{t-1}, y_{t-2}....\}$.

Assumptions

We assume that $E(\upsilon_t \xi'_1) = 0$, $E(w_t \xi'_1) = 0$ and $E(\upsilon_t \omega_\tau) = 0$ for all t and τ. Noting that we can write $\xi_t = \upsilon_t + F\upsilon_{t-1} + F^2\upsilon_{t-2} + .. + F^{t-2}\upsilon_2 + F^{t-1}\xi_1$, we get the following conditions:

a) $E(\upsilon_t \xi'_\tau) = 0$ for all $\tau = t-1, t-2, ...$

b) $E(\omega_t \xi'_\tau) = 0$ for all t and τ.

c) $E(\omega_t y'_\tau) = 0$ for all $\tau = t-1, t-2, ...(\sin ce\ E(\omega_t(A'x_\tau + H'\xi_\tau + \omega_\tau)' = 0)$

d) $E(\upsilon_t y'_\tau) = 0$ for all $\tau = t-1, t-2, ...(\sin ce\ E(\upsilon_t(A'x_\tau + H'\xi_\tau + \omega_\tau)' = 0)$

The Filter

The filter is motivated as an algorithm for calculating linear least squares forecasts of the state vector on the basis of the data observed through date t, $\widehat{\xi}_{t+1|t} = \widehat{E}(\xi_{t+1}|I_t)$, where the operator $\widehat{E}$ denotes the linear projection of ξ_{t+1} *on* I_t and a constant, and $I_t = \left\{y_t', y_{t-1}', y_{t-2}'..y_1', x_t', x_{t-1}', x_{t-2}'..x_1'\right\}'$. The filter calculates these forecasts recursively, generating $\widehat{\xi}_{1|0}$, $\widehat{\xi}_{2|1}$, $\widehat{\xi}_{3|2}$, ...,$\widehat{\xi}_{T|T-1}$ in succession. Associated with each these forecast is a MSE matrix represented by the following (rxr) matrix $P_{t+1|t} = E\left[(\xi_{t+1} - .\widehat{\xi}_{t+1|t})(\xi_{t+1} - .\widehat{\xi}_{t+1|t})'\right]$. For the typical element $\widehat{\xi}_{t|t-1}$, with its associated $P_{t|t-1}$, the goal of the the filter is to produce $\widehat{\xi}_{t+1|t}$, with its associated $P_{t+1|t}$. The steps of the filter typically involve initializing the filter, updating the linear projection (when new information arrives) and producing a new forecast conditional on the new information set.

Initializing the filter: Starting the recursion.

To initialize the filter we need a proxy of $\widehat{\xi}_{1|0}$ and we take for this the unconditional expectation, $E(\xi_1)$, with the associated $P_{1|0} = E\left[(\xi_1 - .E(\xi_1))(\xi_1 - .E(\xi_1))'\right]$. To calculate $E(\xi_1)$, we use the state equation and take expectations in both sides obtaining $E(\xi_{t+1}) = FE(\xi_t)$ or $(I - F)E(\xi_t) = 0$. If all the eigen values of F are smaller than 1 this implies that $\widehat{\xi}_{1|0} = E(\xi_1) = 0$. The associated MSE matrix $P_{1|0} = E\left[(\xi_1)(\xi_1)'\right]$ can be obtained in similar way noting that $E(\xi_{t+1}\xi_{t+1}') = E\left[(F\xi_t + \upsilon_{t+1})(F\xi_t + \upsilon_{t+1})'\right] = FE(\xi_t\xi_t')F' + Q$. If we denote $\Sigma = E(\xi_t\xi_t')$, then we can write the previous expression as $\Sigma = F\Sigma F' + Q$. If all the eigenvalues or F are smaller than 1 then this can be solved using Vec operators as $Vec(P_{1|0}) = Vec(\Sigma) = \left[I_{r^2} - F \otimes F\right]^{-1} Vec(Q)$.

Given the starting values $\widehat{\xi}_{1|0}$ and $P_{1|0}$, the next step is to calculate $\widehat{\xi}_{2|1}$ and $P_{2|1}$.

Forecasting y_t.

To forecast y_t we have to note that we assumed that x_t contains no information about ξ_t beyond that contained in I_{t-1}, then $\widehat{E}(\xi_t|x_t, I_{t-1}) = \widehat{E}(\xi_t|I_{t-1}) = \widehat{\xi}_{t|t-1}$.

Then the forecast of y_t is

$$\widehat{y}_{t|t-1} = A'x_t + H'\widehat{\xi}_{t|t-1},$$

with associated forecasting error $y_t - \widehat{y}_{t|t-1} = H'(\xi_t - \widehat{\xi}_{t|t-1}) + \omega_t$ and MSE, $E\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] = H'P_{t|t-1}H + R$.

Updating the inference about ξ_t.

The inference about the value of ξ_t is updated on the basis of the observation of y_t to produce $\widehat{\xi}_{t|t} = \widehat{E}(\xi_t|y_t, x_t, I_{t-1}) = \widehat{E}(\xi_t|I_t)$.

The formulae to update a linear projection is

$$\widehat{\xi}_{t|t} = \widehat{\xi}_{t|t-1} + E\left[(\xi_t - \widehat{\xi}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right]\left(E\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right]\right)^{-1}(y_t - .\widehat{y}_{t|t-1})$$

Noting that:

-$E\left[(\xi_t - \widehat{\xi}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] = P_{t|t-1}H$,

-$E\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] = H'P_{t|t-1}H + R$,

-$\widehat{y}_{t|t-1} = A'x_t + H'\widehat{\xi}_{t|t-1}$,

this formulae can be written as

$$\widehat{\xi}_{t|t} = \widehat{\xi}_{t|t-1} + P_{t|t-1}H\left(H'P_{t|t-1}H + R\right)^{-1}(\underbrace{y_t - A'x_t + H'\widehat{\xi}_{t|t-1}}_{=H'(\xi_t - \widehat{\xi}_{t|t-1}) + \omega_t})$$

This expression has as associated MSE,

$$E\left[(\xi_t - \widehat{\xi}_{t|t})(\xi_t - \widehat{\xi}_{t|t})'\right] = P_{t|t} = P_{t|t-1} - P_{t|t-1}H\left(H'P_{t|t-1}H + R\right)^{-1}H'P_{t|t-1}$$

Proof $E\left[(\xi_t - \widehat{\xi}_{t|t})(\xi_t - \widehat{\xi}_{t|t})'\right]$

$= E\left[(\xi_t - \left[\widehat{\xi}_{t|t-1} + P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}(y_t - A'x_t + H'\widehat{\xi}_{t|t-1})\right])(\xi_t - \widehat{\xi}_{t|t})'\right]$

$$=E\left[\underbrace{(\xi_t - \widehat{\xi}_{t|t-1})}_{=A} - \underbrace{P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}(H'(\xi_t - \widehat{\xi}_{t|t-1}) + \omega_t)}_{B} (A-B)'\right]$$

$=E(AA') - E(BA') - E(AB') + E(BB')$, *where*

- $E(AA') = P_{t|t-1}$
- $E(BA') = E\left\{\left[P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}(H'(\xi_t - \widehat{\xi}_{t|t-1}) + \omega_t)\right](\xi_t - \widehat{\xi}_{t|t-1})'\right\} = \underline{P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}}.$
- $E(AB') = E\left\{(\xi_t - \widehat{\xi}_{t|t-1})\left[((\xi_t - \widehat{\xi}_{t|t-1})'H + \omega_t')(H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}\right]\right\} = \underline{P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}}$
- $E(BB') = E\left\{\left[P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}(H'(\xi_t - \widehat{\xi}_{t|t-1}) + \omega_t)\right]\left[((\xi_t - \widehat{\xi}_{t|t-1})'H + \omega_t')(H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}\right]\right\}$

$$=P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1} \underbrace{E\left\{\left[(H'(\xi_t - \widehat{\xi}_{t|t-1}) + \omega_t)\right]\left[((\xi_t - \widehat{\xi}_{t|t-1})'H + \omega_t')\right]\right\}}_{=H'P_{t|t-1}H+R} (H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}$$

$=\underline{P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}}.$

Then $E\left[(\xi_t - \widehat{\xi}_{t|t})(\xi_t - \widehat{\xi}_{t|t})'\right] = E(AA') - E(BA') - E(AB') + E(BB')$

$= P_{t|t-1} - P_{t|t-1}H(H'P_{t|t-1}H+R)^{-1}H'P_{t|t-1}.$

Producing a Forecast of ξ_{t+1}

$$\widehat{\xi}_{t+1|t} = \widehat{E}(\xi_{t+1}|I_t) = F\widehat{E}(\xi_t|I_t) + \widehat{E}(\upsilon_{t+1}|I_t) = F\widehat{\xi}_{t|t}.$$

Using the formulae derived for updating a linear projection, we can express this forecast as

$$\widehat{\xi}_{t+1|t} = F\left[\widehat{\xi}_{t|t-1} + P_{t|t-1}H\left(H'P_{t|t-1}H+R\right)^{-1}(y_t - A'x_t + H'\widehat{\xi}_{t|t-1})\right]$$

$$=F\widehat{\xi}_{t|t-1} + \underbrace{FP_{t|t-1}H\left(H'P_{t|t-1}H+R\right)^{-1}}_{=K_t \text{ The Kalman Gain Matrix.}}(y_t - A'x_t + H'\widehat{\xi}_{t|t-1})$$

The MSE Associated with the forecast can easily be obtained from the forecasting equation

$$\begin{aligned}
P_{t+1|t} &= E\left[(F\xi_t + \upsilon_{t+1} - .F\widehat{\xi}_{t|t})(F\xi_t + \upsilon_{t+1} - F\widehat{\xi}_{t|t})'\right] \\
&= FP_{t|t}F' + Q \\
&= F\left[P_{t|t-1} - P_{t|t-1}H\left(H'P_{t|t-1}H+R\right)^{-1}H'P_{t|t-1}\right]F' + Q
\end{aligned}$$

Examples of State Representation.

$$\begin{aligned} \xi_{t+1} &= F\xi_t + \upsilon_{t+1} && \textit{State Equation} \\ y_t &= A'x_t + H'\xi_t + \omega_t && \textit{Observation Equation} \end{aligned}$$

1. An ARMA Process

   Consider the following ARMA Process

$$\begin{aligned} (y_{t+1}-\mu) &= \phi_1(y_t-\mu)+\phi_2(y_{t-1}-\mu)+...+\phi_P(y_{t-p+1}-\mu)+\varepsilon_{t+1} \\ E(\varepsilon_t\varepsilon_\tau) &= \begin{cases} \sigma^2 \text{ if } t=\tau \\ 0 \text{ Otherwise.} \end{cases} \end{aligned}$$

   The state equation

$$\underbrace{\begin{bmatrix} y_{t+1}-\mu \\ y_t-\mu \\ \\ \\ y_{t-p+2}-\mu \end{bmatrix}}_{=\xi_{t+1}} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & & \phi_P \\ 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \\ & & 1 & 0 \end{bmatrix}}_{=F} \underbrace{\begin{bmatrix} y_t-\mu \\ y_{t-1}-\mu \\ \\ \\ y_{t-p+1}-\mu \end{bmatrix}}_{=\xi_t} + \underbrace{\begin{bmatrix} \varepsilon_{t+1} \\ 0 \\ \\ \\ 0 \end{bmatrix}}_{=\upsilon_{t+1}}$$

   Observation Equation (Identity)

$$\begin{aligned} y_t &= \mu + \begin{bmatrix} 1, & 0, & & 0 \end{bmatrix} \begin{bmatrix} y_t-\mu \\ y_{t-1}-\mu \\ \\ \\ y_{t-p+1}-\mu \end{bmatrix} \\ \textit{where } A' &= \mu,\ x_t=1,\ H'=\begin{bmatrix} 1, & 0, & & 0 \end{bmatrix},\ \omega_t=0. \end{aligned}$$

2. A $MA(1)$ process

   Consider the following MA process

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$

   The state Equation

$$\begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_t \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \end{bmatrix}$$

The Observation Equation (Identity)

$$
\begin{aligned}
y_t &= \mu + \begin{bmatrix} 1 & \theta \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{bmatrix} \\
where\ A' &= \mu,\ x_t = 1,\ H' = \begin{bmatrix} 1, & \theta \end{bmatrix},\ \omega_t = 0.
\end{aligned}
$$

3. An $ARMA(p,q)$

Consider the following ARMA Process

$$
\begin{aligned}
(y_t - \mu) &= \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + ... + \phi_P(y_{t-r} - \mu) + \varepsilon_{t+1} \\
&\quad + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + + \theta_2 \varepsilon_{t-r+1} \\
E(\varepsilon_t \varepsilon_\tau) &= \begin{cases} \sigma^2 \text{ if } t = \tau \\ 0 \text{ Otherwise.} \end{cases} \quad \text{and } r = Max\{p, q+1\}, \\
\phi_j &= 0 \text{ for } j > p \text{ and } \theta_j = 0 \text{ for } j > q.
\end{aligned}
$$

The state equation

$$
\underbrace{\begin{bmatrix} y_{t+1} - \mu \\ y_t - \mu \\ \\ \\ y_{t-p+2} - \mu \end{bmatrix}}_{=\xi_{t+1}} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & & \phi_P \\ 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ & & & \\ & & 1 & 0 \end{bmatrix}}_{=F} \underbrace{\begin{bmatrix} y_t - \mu \\ y_{t-1} - \mu \\ \\ \\ y_{t-p+1} - \mu \end{bmatrix}}_{=\xi_t} + \underbrace{\begin{bmatrix} \varepsilon_{t+1} \\ 0 \\ \\ \\ 0 \end{bmatrix}}_{=\upsilon_{t+1}}
$$

Observation Equation (Identity)

$$
\begin{aligned}
y_t &= \mu + \begin{bmatrix} 1, & \theta_1, & & \theta_{r-1} \end{bmatrix} \underbrace{\begin{bmatrix} y_t - \mu \\ y_{t-1} - \mu \\ \\ \\ y_{t-p+1} - \mu \end{bmatrix}}_{=\xi_t} \\
where\ A' &= \mu,\ x_t = 1,\ H' = \begin{bmatrix} 1, & \theta_1, & & \theta_{r-1} \end{bmatrix},\ \omega_t = 0.
\end{aligned}
$$

Proof. Notice that the second element of ξ_{t+1} is equal to the first element of ξ_t. We denote this relationship as $\xi_{t+1}[2] = \xi_t[1]$. Then is easy to see that $\xi_{t+1}[2] = \xi_t[1] = L\xi_{t+1}[1]$, $\xi_{t+1}[3] = \xi_t[2] = L^2\xi_{t+1}[1]$, or in general $\xi_{t+1}[j] = \xi_t[j+1] = L^{j-1}\xi_{t+1}[1]$.

Then using the first row of the state equation we can write

$$
\xi_{t+1}[1] = (\phi_1 + \phi_2 L + ... + \phi_r L^{r-1})\xi_t[1] + \varepsilon_{t+1}
$$

or

$$(1+\phi_1 L+\phi_2 L^2+.....+\phi_r L^r)\xi_{t+1}[1]=\varepsilon_{t+1}$$

Using the observation equation we may see that

$$y_t=\mu+(1+\theta_1 L+\theta_2 L^2+.....+\theta_{r-1}L^{r-1})\xi_t[1]$$

Then multiplying the observation equation in both side by $(1+\phi_1 L+\phi_2 L^2+ ..... +\phi_r L^r)$ we obtain

$$\begin{aligned}(1+\phi_1 L+..+\phi_r L^r)(y_t-\mu) &= (1+\theta_1 L+..+\theta_{r-1}L^{r-1})\underbrace{(1+\phi_1 L+..+\phi_r L^r)\xi_t[1]}_{=\varepsilon_t} \\ &= (1+\theta_1 L+..+\theta_{r-1}L^{r-1})\varepsilon_t\end{aligned}$$

■

4. The ex-ante real interest rate

   The ex-ante real interest rate is unobserved. Assume we can write the de-meaned real interest rate as $\xi_t=i_t-\pi^e-\mu$, then we can write the state equation as $\xi_{t+1}=\phi\xi_t+\upsilon_{t+1}$. The econometrician has observations on the ex-post real rate which can be written as

   $$i_t-\pi=(i_t-\pi^e)+(\pi^e-\pi).$$

   Now, if expectations are rational, then $\pi=\pi^e+\omega_t$,. and then the measurement equation is of the form

   $$\begin{aligned}i_t-\pi &= (i_t-\pi^e)+\omega_t \\ &= \mu+\xi_t+\omega_t.\end{aligned}$$

5. Uncovering the cyclical component

   Stock and Watson (1991) postulated the existence of an unobserved variable C_t which represents the state of the business cycle. They assumed that we observed n macroeconomic variables and that these variables, $(y_{1t}, y_{2t}, y_{3t}, ..., y_{nt})$ are assumed to be influenced by the business cycle and also have an idiosyncratic component, denoted X_{it} unrelated to the movements in y_{jt} for $i\neq j$.

   If the business cycle and each of the idiosyncratic components could be described by an univariate AR(1) process, then we can write the measurement equation as

   $$\underbrace{\begin{bmatrix} C_{t+1} \\ X_{1_{t+1}} \\ X_{2_{t+1}} \\ \\ X_{n_{t+1}} \end{bmatrix}}_{=\xi_{t+1}} = \underbrace{\begin{bmatrix} \phi_C & 0 & 0 & 0 & 0 \\ & \phi_1 & & 0 & 0 \\ & & \phi_2 & 0 & 0 \\ & & & . & 0 \\ & & & & \phi_n \end{bmatrix}}_{=F} \underbrace{\begin{bmatrix} C_t \\ X_{1_t} \\ X_{2_t} \\ \\ X_{n_t} \end{bmatrix}}_{=\xi_t} + \underbrace{\begin{bmatrix} \upsilon_{C_{t+1}} \\ \upsilon_{1_{t+1}} \\ \upsilon_{2_{t+1}} \\ \\ \upsilon_{n_{t+1}} \end{bmatrix}}_{=\upsilon_{t+1}}$$

The observation equation can be written as

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ \\ y_{nt} \end{bmatrix}_{nx1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \\ \mu_n \end{bmatrix}_{nx1} + \begin{bmatrix} \gamma_1 & 1 & 0 & 0 & 0 & 0 \\ \gamma_2 & 0 & 1 & 0 & & 0 \\ & & 0 & 0 & & 0 \\ & & & . & & 0 \\ \gamma_n & & & & 0 & 1 \end{bmatrix}_{nx(n+1)} \begin{bmatrix} C_t \\ X_{1_t} \\ X_{2_t} \\ \\ X_{n_t} \end{bmatrix}_{(n+1)x1}$$

6. Linear Regression Models with Time Varying Coefficients

One important application of the state space model with stochastically varying parameters as a regression in which the coefficient vector changes over time.

Consider the following regression model with time varying coefficients.

$$y_t = x_t' \beta_t + \omega_t,$$

in the state space representation this equation represents the measurement equation while we can write the state equation as

$$\underbrace{\left(\beta_{t+1} - \overline{\beta}\right)}_{=\xi_{t+1}} = F \underbrace{\left(\beta_t - \overline{\beta}\right)}_{=\xi_t} + \upsilon_{t+1}$$

If the eigenvalues of F are all inside the unit circle, then $\overline{\beta}$ has the interpretation of the average of the steady-state value for the coefficient vector. If

$$\begin{matrix} \nu_{t+1} \\ \omega_t \end{matrix} \Bigg| x_t,\ I_{t-1} \tilde{} N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0' & \sigma^2 \end{bmatrix} \right),$$

then the state space model can be written as

$$\begin{aligned} y_t &= x_t' \overline{\beta} + x_t' \xi_t + \omega_t, \\ \left(\beta_{t+1} - \overline{\beta}\right) &= F \left(\beta_t - \overline{\beta}\right) + \upsilon_{t+1} \end{aligned}$$