# Handout 11

# 1 Overview of the Kalman Filter

A good reference to pick up the basics of the Kalman filter is the chapter on the Kalman filter in Hamilton's time series volume. A more in depth treatment of the Kalman filter can be found in Durbin and Koopman.

The state space representation used in the setup of the Kalman filter includes two components: a state equation, which describes the evolution of the state variables, and an observation equation, which relates the state variables to the observed variables.

State equation

$$\xi_{t+1} = F\xi_t + v_{t+1}, \quad \text{where } E(v_{t+1}v_{t+1}') = Q.$$

Assume that $v_t$ is multivariate normal.

Observation equation

$$y_t = H'\xi_t.$$

## 1.1 Linear Least Squares Forecasts

Define $\hat{\xi}_{t+1}$ as the projection of $\xi_{t+1}$ on the history of $y_t$.

Define $P_{t+1|t}$ to be the mean squared error matrix associated with the projection

$P_{t+1|t} = E[(\xi_{t+1} - \hat{\xi}_{t+1|t})(\xi_{t+1} - \hat{\xi}_{t+1|t})']$

The Kalman filter gives us a way to recursively calculate $\hat{\xi}_{t+1|t}$.

$$\hat{\xi}_{t+1|t} = F\hat{\xi}_{t|t-1} + K_t(y_t - H'\hat{\xi}_{t|t-1})$$

where $K_t$ is the Kalman gain matrix

$$K_t = FP_{t|t-1}H(H'P_{t|t-1}H)^{-1}$$

In turn, the mean square error matrix is updated following

$$P_{t+1|t} = (F - K_tH')P_{t|t-1}(F' - HK'_t) + Q$$

## Steady-State Kalman Filter

If the eigenvalues of $F$ are all within the unit circle and $Q$ is strictly positive definite then

$P_{t+1/t}$ will converge to $P$.

$K_{t+1}$ will converge to K.

Thus the Kalman filter recursion becomes

$$\hat{\xi}_{t+1|t} = F\hat{\xi}_{t|t-1} + K(y_t - H'\hat{\xi}_{t|t-1})$$

This fact can be used to speed the recursion when dealing with large matrices and large datasets.

## 1.2   Likelihood Function

If all the state variables were observed, we could form the likelihood using

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

but strategy fails when $\xi_t$ is not directly observed.

Instead use the fact that

$$a_t = K_t^{-1}(\hat{\xi}_{t+1|t} - F\hat{\xi}_{t|t-1})$$

where $a_t = (y_t - H'\hat{\xi}_{t|t-1})$ is the forecast error for the observed variables.

Notice that $a_t$ is uncorrelated over time by construction. If $v_t$ is multivariate normal, then $a_t$ is also multivariate normal.

The var/cov matrix of $a_t$ is given by $H'P_{t|t-1}H$.

$$g(a_t) = (2\pi)^{-\frac{n}{2}} \left| H'P_{t|t-1}H \right|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}[a_t'(H'P_{t|t-1}H)^{-1}a_t] \right\} \tag{1}$$

## 1.3   Initialization schemes

To form the unconditional likelihood, set $\hat{\xi}_{1|0} = E\xi_t$ and set $P_{1|0}$ to unconditional variance covariance of $\xi_t$, where the variance is

$$E\left[\xi_t \xi_t'\right] = E\left[(F\xi_{t-1} + v_t)(F\xi_{t-1} + v_t)'\right],$$

which can be written as:

$$\Sigma = F\Sigma F' + Q.$$

See Hamilton on how to use the properties of the vec operator to solve the equation above for *Sigma*.

This strategy has the drawback that $\xi = 0$ might imply structural innovations so outlandish that they have a disproportionate weight in the likelihood when the sample is small. Another drawback is that this scheme is viable only when $\xi_t$ is stationary.

## 1.4   Fixed and known initial point

Set $\xi_0 = 0$.

Set $P_0 = 0$.

To avoid polluting the likelihood, train the filter by running the updates on a pre-sample, but not incorporating the information into the likelihood.

## 1.5    The augmented Kalman filter

let $\xi_0 = z_0$

Consider two possibilities:

1) let $var(z_0) = kI$ where $k \to \infty$ (Rosenberg)

2) let $z_0$ be non-stochastic but unknown (de Jong), estimate jointly with other parameters

These schemes also work when some of the state variables are trending.

## 1.6    Filtering and Smoothing

The history of $\hat{\xi}_{t|t-1}$ is referred to as the filtered estimates of the state variables. Notice that the information set used in forming these estimates varies through time, and incorporates new information as we go along.

The Kalman filter can also help us produce estimates of history of the state variables conditional on the information set implied by all of the observations available. The history of the state variables conditional on all of the observations is referred to as the smoothed estimates of the state variables.

The setup for the Kalman filter in Hamilton's "Time Series Analysis" assumes that we the state equation has as many shocks as state variables. The smoothing recursion derived in Hamilton runs into a statistical singularity that can be avoided. Durbin and Koopman show us how to deal with generalizing the the smoothing recursion to cases state equations that have fewer shocks than state variables. Below is a cheat sheet that matches relates the notation used in Hamilton and in Durbin-Koopman.

Kalman filter setup using the notation in Durbin-Koopman

$$\begin{aligned} \alpha_{t+1} &= T\alpha_t + R\eta_t \\ y_t &= Z\alpha_t + \varepsilon_t \end{aligned}$$

with $\varepsilon_t = N(0, H)$, $\eta = N(0, Q)$.

Kalman setup using the notation in Hamilton

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

$$y_t = A'x_t + H'\xi_t + w_t$$

with $w_t = N(0, R)$, $v_t = (0, Q)$.

| Durbin- Koopman | Hamilton |
|---|---|
| $T$ | $F$ |
| $RQR'$ | $Q$ |
| $Z$ | $H'$ |
| $H$ | $R$ |
| $\alpha_{t+1}$ | $\xi_{t+1}$ |
| $a_{t+1}$ | $\hat{\xi}_{t+1|t}$ |
| $\hat{\alpha}_t$ | $\hat{\xi}_{t|T}$ |
| $K_t$ | $K_t$ |
| $P_t$ | $P_{t|t-1}$ |
| $F_t$ | $H'P_{t|t-1}H + R$ |

We want to translate

$$r_{t-1} = Z'F_t^{-1}v_t + L_t'r_t$$

$$\hat{\alpha}_t = a_t + P_t r_{t-1}$$

where

$$F_t = ZP_tZ' + H$$

$$L_t = T - K_tZ$$

$$K_t = TP_tZ'F_t^{-1}$$

$$P_{t+1} = TP_tL_t' + RQR'$$

$$v_t = y_t - Za_t$$

$$a_{t+1} = Ta_t + K_tv_t$$

Start from from Durbin-Koopman's

$$r_{t-1} = Z'F_t^{-1}v_t + L_t'r_t.$$

Using Hamilton's notation, this translates to:

$$r_{t-1} = H\left(H'P_{t|t-1}H + R\right)^{-1}\left(y_t - H'\hat{\xi}_{t|t-1}\right) + (F' - HK_t')r_t$$

and Durbin-Koopman's updating equation

$$\hat{\alpha}_t = a_t + P_t r_{t-1}$$

can be re-written in Hamilton's notation as

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t-1} + P_{t|t-1}r_{t-1}$$

How to get $r_t$? Start from the end. By construction $r_T$ is equatl to 0.

$$r_{T-1} = H\left(H'P_{T|T-1}H + R\right)^{-1}\left(y_T - H'\hat{\xi}_{T|T-1}\right) + (F' - HK_T')r_T$$

Substituting the above into

$$\hat{\xi}_{T|T} = \hat{\xi}_{T|T-1} + P_{T|T-1}r_{T-1}$$

results in:

$$\hat{\xi}_{T|T} = \hat{\xi}_{T|T-1} + P_{T|T-1}H\left(H'P_{T|T-1}H + R\right)^{-1}\left(y_T - H'\hat{\xi}_{T|T-1}\right)$$

but we know $\hat{\xi}_{T|T-1}$ and $P_{T|T-1}$ from the filtering recursions. So now we can iterate backwards to get $\hat{\xi}_{T-1|T}$, $\hat{\xi}_{T-2|T}$, and so on.

From Durbin-Koopman, we also learn how to obtain the smoothed estimates for the shocks in the state equation (in their notation):

$$\hat{\eta}_t = Q_t R_t' r_t.$$