A (very) simple Kalman Filter : Application to inflation

Brendan Berthold

4/8/2020

Introduction

What is the Kalman Filter and how it is used in economics.

This implementation is interesting because:

- it provides a step by step introduction to an extremely simple Kalman filter
- it aims to give you the basic intuition behind the Kalman filter with a very simple example
- everything is in one dimension to avoid potentially complex matrix otpimisation
- the idea is to focus only on the intuition and the basic implementation in R
- does not cover the maximum likelihood estimation of the parameters (once again in an attempt to simplify as much as possible) but could be the subject of another document if there is demand
- illustrates the Kalman filter in an economic context with an application the Swiss inflation rate

The theory

Notation

The general form of the Kalman filter as presented in Hamilton Chapter 13 (add ref) is given by a "measurement equation":

$$y_t = A'x_t + H'\xi_t + w_t \tag{1}$$

With $E(w_t w_t') = R$

And a transition (or state) equation:

$$\xi_t = F\xi_{t-1} + v_t \tag{2}$$

With $E(v_t v_t') = Q$.

Notation:

- y_t is the vector of observed variables (i.e. the data)
- x_t is a vector of deterministic components (we won't spend time on it in this document)
- ξ_t is the unobserved "state" variables
- w_t and v_t are unobserved, mutually and serially uncorrelated noise variables
- A, H, R, F, and Q are non-random "system" variables matrices that may depend on unknown parameters (some of them can be retrieved using Maximum Likelihood estimations)

The general system defined by (1) and (2) is flexible and can accomodate a variety of representation. For instance, a standard AR(p) process fits into the general notation in the following way:

Let $y_t \sim AR(p)$, that is:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \epsilon_t$$

This process can be represented as a "state-space" model in the following way:

$$
\xi_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}
$$

$$
F = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_{p-1} & \phi_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 1 & 0 \end{bmatrix}
$$

$$
v_t = \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

And $w_t = 0, A = 0$, and $H' = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}$

Procedure and idea of the Kalman Filter

Notation:

- $y_{1:t} = \{y_i\}_{i=1}^t$
- $\xi_{t|k} = E(\xi_t | y_{1:k})$
- $P_{t|k} = Var(\xi_t | y_{1:k})$

In words, the Kalman filter is a recursive algorithm to construct $\xi_{t|t}$ and $P_{t|t}$ from known values in t, that is $y_t, x_t, \xi_{t-1|t-1}, P_{t-1|t-1}$.

To derive the filter, we assume that both w_t and v_t follow iid Gaussian process, that is:

$$
\begin{bmatrix} w_t \\ v_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix} \right)
$$

This notably implies that both y_t and ξ_t follow a *joint* Normal distribution. In that context, the best estimator (in the sense that it minimises the mean squared error) is given by the conditional expectation.

To find the conditional expectation of ξ_t and y_t (that is $\xi_{t|t}$ and $y_{t|t}$), we can use the following theorem on the conditional distribution of a multivariate normal:

Suppose that:

$$
\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)
$$

Then:

$$
\begin{aligned}
E(z_1 | z_2) &= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(z_2 - \mu_2) \\
Var(z_1 | z_2) &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}
\end{aligned}
$$

This theorem is the *key idea* of the Kalman Filter. In particular, the Kalman Filter is "simply" an application of it.

Defining $z_1 = \xi_t$ and $z_2 = y_t$, and recognizing that ξ_t and y_t are jointly Normal conditional on past values, we can write the following:

$$
\begin{bmatrix} \xi_t \\ y_t \end{bmatrix} \Bigg| y_{1:t-1} \sim \mathcal{N}\left(\begin{bmatrix} \xi_{t|t-1} \\ y_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & \Sigma_{\xi,y|t-1} \\ \Sigma_{\xi,y|t-1} & \Sigma_{yy|t-1} \end{bmatrix} \right)
$$

Using the formula of the conditional normal:

$$\xi_{t|t} = \xi_{t|t-1} + \Sigma_{\xi,y|t-1}\Sigma^{-1}_{yy|t-1}(y_t - y_{t|t-1})$$
$$P_{t|t} = P_{t|t-1} - \Sigma_{\xi,y|t-1}\Sigma^{-1}_{yy|t-1}\Sigma_{\xi,y|t-1}$$

1. $\xi_{t|t-1} = F\xi_{t-1|t-1}$
2. $y_{t|t-1} = A'x_t + H'\xi_{t|t-1}$
3. $P_{t|t-1} = FP_{t-1|t-1}F' + Q$
4. $\Sigma_{yy|t-1} = H'P_{t|t-1}H + R \equiv h_t$
5. $\Sigma_{\xi,y|t-1}\Sigma^{-1}_{yy|t-1} = P_{t|t-1}Hh_t^{-1} \equiv K_t$
6. $\eta_t = y_t - y_{t|t-1}$

Thus:

$$\xi_{t|t} = \xi_{t|t-1} + K_t\eta_t$$
$$P_{t|t} = P_{t|t-1} - K_tH'P_{t|t-1}$$

This procedure allows us to retrieve $\xi_{t|t}$ and $P_{t|t}$ recursively (i.e. assuming $\xi_{t-1|t-1}$ and $P_{t-1|t-1}$) are known.

Application

To better understand the algorithm let us consider the following (uni-dimensional) simple example. For simplicity, we assume that ϕ is known but it could also be retrieved using a MLE approach.

The state space model is of the form:

$$y_t = \phi\xi_t + w_t$$
$$\tau_t = \tau_{t-1} + v_t$$

The sample variance and covariance (see R Script for computations) are:

```
var_Y_t
```

```
## [1] 0.7712221
```

```
cov_Y_1
```

```
## [1] -0.2622073
```

The first step is to approximate the value of R and Q, that is the variance of w_t and v_t respectively.

To do so, we can recognize that:

$$\Delta Y_t = \Delta\tau_t + \Delta\epsilon_t$$
$$= \eta_t + \Delta\epsilon_t$$

Using the fact that ϵ and η are independente and covariance stationary, the variance is given by:

$$Var(\Delta Y_t) = \sigma^2_\eta + 2\sigma^2_\epsilon$$

Similarly, we can compute:

$$Cov(\Delta Y_t, \Delta Y_{t-1}) = -\sigma^2_\epsilon$$

Which implies:

```r
sigma_sq_e <- -cov_Y_1
sigma_sq_e
```

```
## [1] 0.2622073
```

```r
sigma_sq_eta <- var_Y_t - 2*sigma_sq_e
sigma_sq_eta
```

```
## [1] 0.2468075
```

To approximate the variance of τ_0, we can use that:

$$Var(\tau_1) = Var(\tau_0) + \sigma_\epsilon^2$$

We can now use the conditional variance formula of an AR(1) process:

$$Var(\tau_t) = t\sigma_\eta^2$$

We also assume $\tau_{0|0} = 0$

Kalman filter

Since we know $E(\tau_0|t=0) = \tau_{0|0}$ $Var(\tau_t|t=0) = P_{0|0}$, we can recursively compute $\tau_{t|t}$ and $P_{t|t}$ for $t > 0$ using the Kalman Filter procedure.

1. $\tau_{t|t-1} = \tau_{t-1|t-1}$
2. $Y_{t|t-1} = \tau_{t|t-1}$
3. $P_{t|t-1} = P_{t-1|t-1} + \sigma_\eta^2$
4. $Var(Y_t|t-1) = h_t = P_{t|t-1} + \sigma_\epsilon^2$
5. $Cov(\tau_t, Y_t|t-1) \times h_t = K_t = P_{t|t-1} \times h_t^{-1}$
6. $y_t - y_{t|t-1}$

Using this, we get our next period forecast:

1. $\tau_{t|t} = \tau_{t-1|t} + K_t(y_t - y_{t|t-1}$
2. $P_{t|t} = P_{t|t-1} - K_t \cdot Cov(\tau_t, Y_t|t-1)$

aparté : Likelihood function

Let consider the following model:

$$\begin{aligned} Y_t &= \tau_t + \epsilon_t \\ \tau_t &= \phi\tau_{t-1} + \eta_t \end{aligned}$$

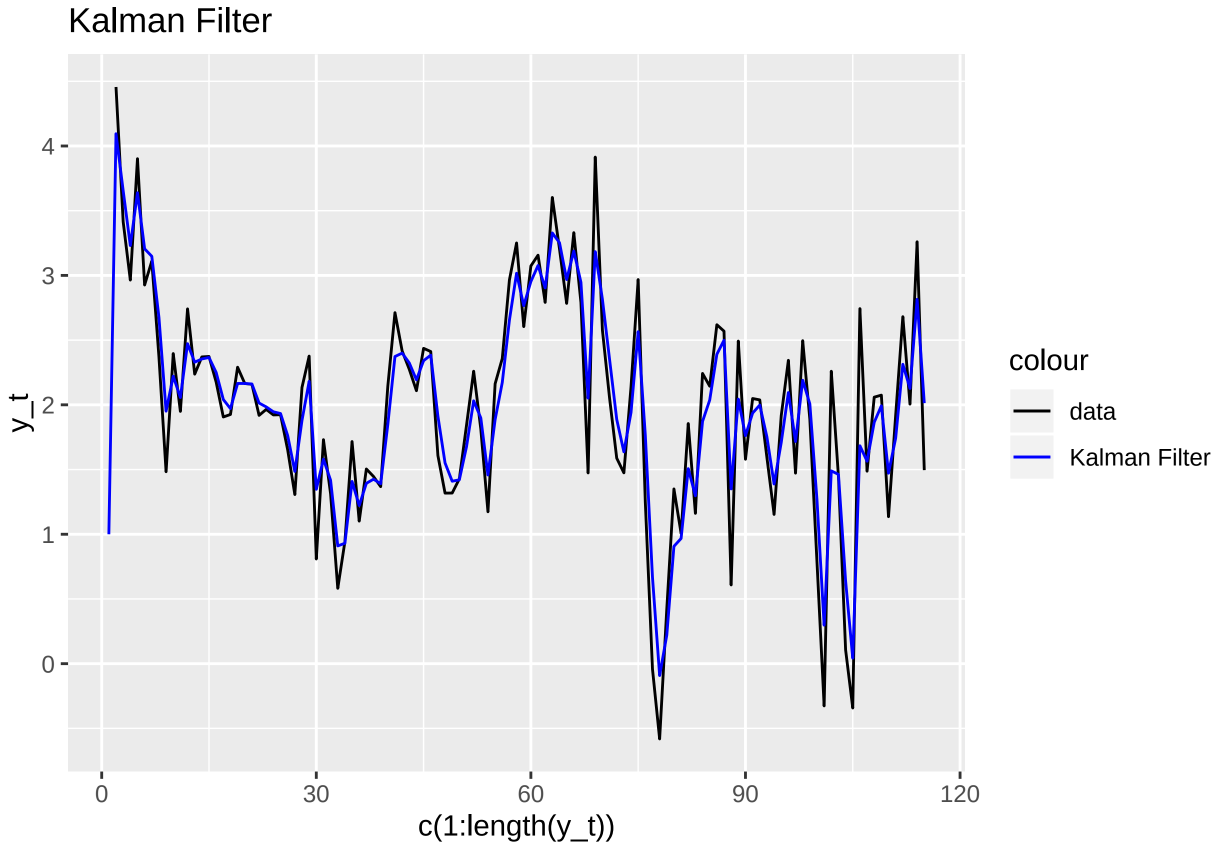
We can deduce:

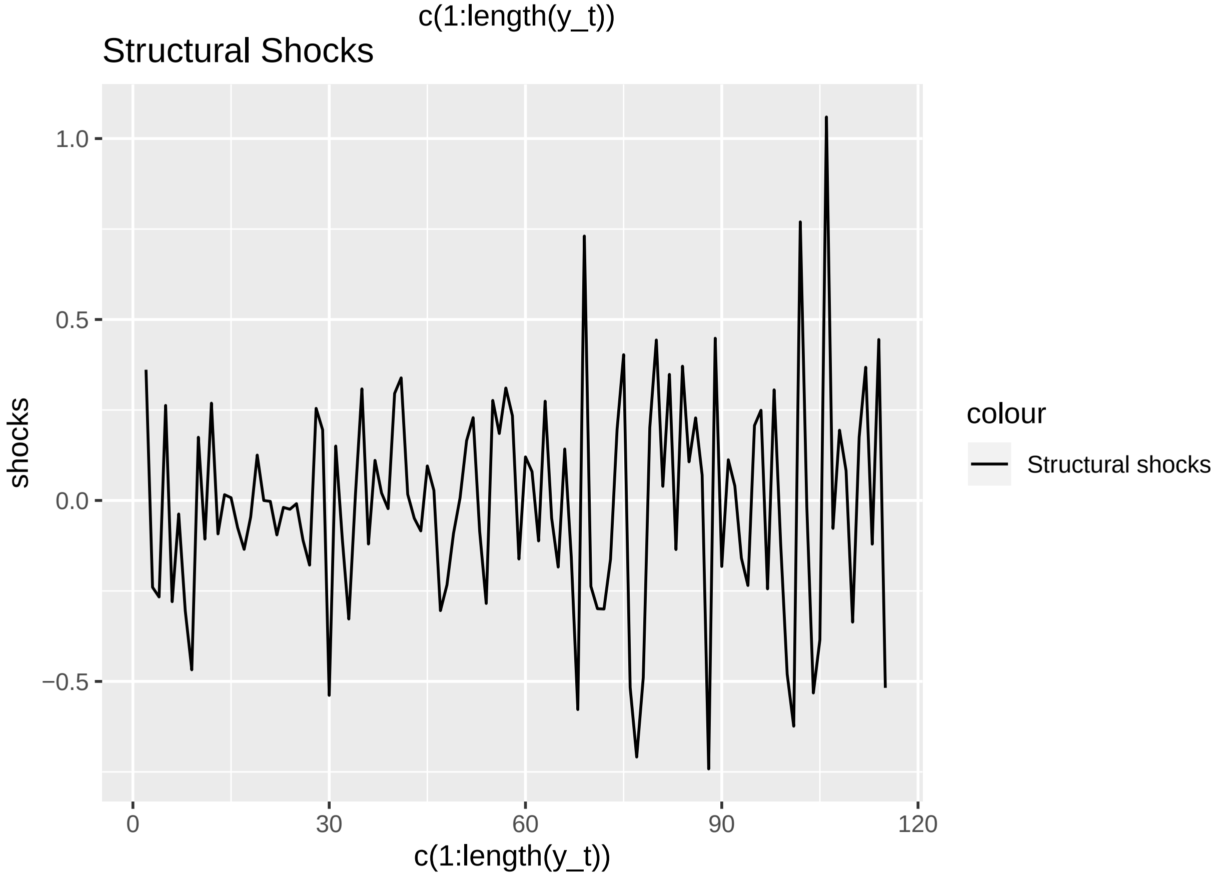
$$\tau_{t|t-1} \sim N(\phi\tau_{t-1|t-1}, V = \phi^2 P_{t-1|t-1} + \sigma_\eta^2)$$

Which implies the conditional pdf is given by:

$$f(\tau_t|t-1) = \frac{1}{\sqrt{2\phi V}} e^{-\frac{1}{2}\left(\frac{Y_t - \phi\tau_{t-1|t-1}}{V}\right)^2}$$

Recursive computation in R





As can be seen from the first two steps of the Kalman filter, the best forecast for next period (that is 2018:Q4) inflation today (that is 2018:Q3) is equal to $\tau_{2018:Q4|2018Q3}$.

According to the code it is:

```
l.tau[length(l.tau)]
```

```
## [1] 2.815144
```

With variance given by:

```
l.P[length(l.P)]
```

```
## [1] 0.4061461
```

References

Hamilton Mark Watson's courses in Gerzensee