Econ 8185 (002): Quant PS1

Bipul Verma

November 8, 2021

Abstract

This document applies Kalman Filter, EM algorithm to US gdp data.

1 US GDP growth rate

The GDP growth rate is given by the following formula:

$$g_t = \frac{Y_{t+1} - Y_t}{Y_t}$$

Another equivalent way to calculate the growth rate is to take differences of logs. This follows since:

$$\begin{aligned} log(Y_{t+1}) - log(Y_t) &= log(\frac{Y_{t+1}}{Y_t}) \\ &= log(1 + \frac{Y_{t+1}}{Y_t} - 1) \\ &= log(1 + g_t) \\ &\approx g_t. \end{aligned}$$

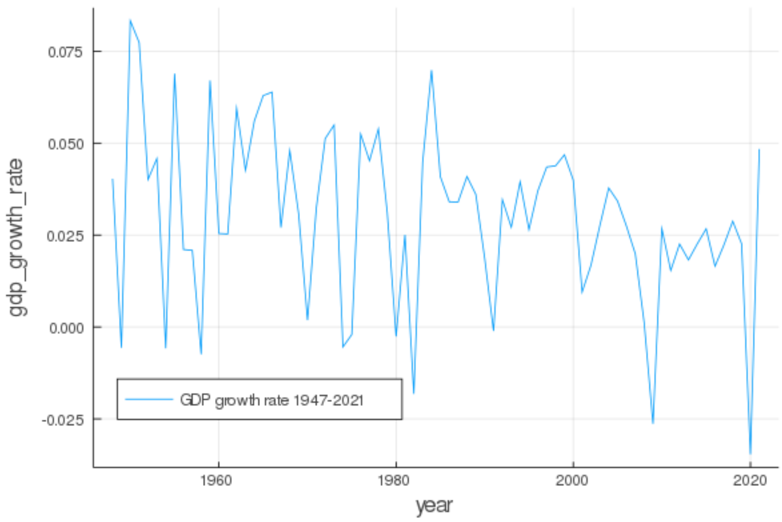

Figure 1: US GDP growth rate 1947-2021

Summary of Kalman Filter, Smoother, and EM Algorithm

Consider the following state space system:

$$\begin{aligned} x_t &= Bx_{t-1} + w_t \\ y_t &= Zx_t + v_t \\ w_t &\sim N(0, Q) \\ v_t &\sim N(0, R). \end{aligned}$$

1.1 Kalman Filter

Kalman filter gives the prediction of state and its update based on past value of observable.

Notations:

$$\hat{x}_{t-1|t-1} = \mathbb{E}[x_{t-1}|y^{t-1}]$$
$$P_{t-1|t-1} = \mathbb{E}[(x_{t-1} - \hat{x}_{t-1|t-1})(x_{t-1} - \hat{x}_{t-1|t-1})'|y^{t-1}]$$
$$\hat{x}_{t|t-1} = \mathbb{E}[x_t|y^{t-1}]$$
$$P_{t|t-1} = \mathbb{E}[(x_t - \hat{x}_{t|t-1})(x_t - \hat{x}_{t|t-1})'|y^{t-1}]$$

Initialization: Set $\hat{x}_{1|0} = x_0$ and $P_{1|0} = P_0$.
Filtering Algorithm (Forward Pass) $t \geq 1$:

$$\kappa_t = P_{t|t-1}Z'(ZP_{t|t-1}Z' + R)^{-1}$$
$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t(y_t - Z\hat{x}_{t|t-1})$$
$$P_{t|t} = P_{t|t-1} - \kappa_t ZP_{t|t-1}$$
$$\hat{x}_{t+1|t} = B\hat{x}_{t|t}$$
$$P_{t+1|t} = BP_{t|t}B' + Q$$

1.2 Kalman Smoother:

The estimates of states from the Kalman Filter will feed into the Kalman Smoother to obtain estimates based on all values of observable.
Notations:

$$\hat{x}_{t|T} = \mathbb{E}[x_t|Y]$$

Smoothing Algorithm (Backward Pass):

$$J_t = P_{t|t}B'P_{t+1|t}^{-1}$$
$$\hat{x}_{t|T} = \hat{x}_{t|t} + J_t(\hat{x}_{t+1|T} - \hat{x}_{t+1|t})$$
$$P_{t|T} = P_{t|t} + J_t(P_{t+1|T} - P_{t+1|t})J_t'$$
$$P_{t,t-1|T} = J_{t-1}P_{t|T}$$

1.3 EM Algorithm

For the purpose of current exercise we'll write down the update equation for Q and R only. Note that the expectations are conditional on (Y, θ_j).
Parameter Update Equation:

$$Q_{j+1} = \frac{1}{T-1}\sum_{t=2}^{T}(\mathbb{E}[x_t x_t'] - \mathbb{E}[x_t x_{t-1}']B' - B\mathbb{E}[x_{t-1}x_t'] + B\mathbb{E}[x_{t-1}x_{t-1}']B')$$
$$R_{j+1} = \frac{1}{T}\sum_{t=1}^{T}(\mathbb{E}[y_t y_t'] - \mathbb{E}[y_t x_t']Z' - Z\mathbb{E}[x_t y_t'] + Z\mathbb{E}[x_t x_t']Z')$$

The update equations are sometimes written in the following form as well:

$$Q_{j+1} = \frac{1}{T-1}\left(\sum_{t=2}^{T}\mathbb{E}[x_t x_t'] - \left(\sum_{t=2}^{T}\mathbb{E}[x_t x_{t-1}']\right)\left(\sum_{t=2}^{T}\mathbb{E}[x_{t-1}x_{t-1}']\right)^{-1}\left(\sum_{t=2}^{T}\mathbb{E}[x_{t-1}x_t']\right)\right)$$
$$R_{j+1} = \frac{1}{T}\left(\sum_{t=1}^{T}\mathbb{E}[y_t y_t'] - \left(\sum_{t=1}^{T}\mathbb{E}[y_t x_t']\right)\left(\sum_{t=1}^{T}\mathbb{E}[x_t x_t']\right)^{-1}\left(\sum_{t=1}^{T}\mathbb{E}[x_t y_t']\right)\right).$$

The conditional expectations in the update equation can be evaluated from the Kalman Smoother output as follows:

$$\mathbb{E}[x_t x_t'|Y,\theta_j] = P_{t|T} + \hat{x}_{t|T}\hat{x}_{t|T}'$$
$$\mathbb{E}[x_t x_{t-1}'|Y,\theta_j] = P_{t,t-1|T} + \hat{x}_{t|T}\hat{x}_{t-1|T}'$$

1.3.1 Implementation of EM Algorithm

To understand the implementation we introduce some additional notations for clarity.
Notations:

$$\begin{aligned}
\alpha_t &= \mathbb{E}[y_t y_t'|Y,\theta_j] \\
\delta_t &= \mathbb{E}[y_t x_t'|Y,\theta_j] \\
\gamma_t &= \mathbb{E}[x_t x_t'|Y,\theta_j] \\
\beta_t &= \mathbb{E}[x_t x_{t-1}'|Y,\theta_j].
\end{aligned}$$

All the above mentioned expectations can be calculated from the output of Kalman Smoother as outlined in the previous section. The update equations using the above notation is given below.
Update Equation:

$$\begin{aligned}
Q_{j+1} &= \frac{1}{T-1}\sum_{t=2}^{T}(\gamma_t - \beta_t B' - B\beta_t' + B\gamma_{t-1}B') \\
R_{j+1} &= \frac{1}{T}\sum_{t=1}^{T}(\alpha_t - \delta_t Z' - Z\delta_t' + Z\gamma_t Z')
\end{aligned}$$

The update equations are sometimes written in the following form as well. Writing the update equation in the below form requires making use of some trace manipulations.

$$\begin{aligned}
Q_{j+1} &= \frac{1}{T-1}\left(\sum_{t=2}^{T}\gamma_t - \Big(\sum_{t=2}^{T}\beta_t\Big)\Big(\sum_{t=2}^{T}\gamma_{t-1}\Big)^{-1}\Big(\sum_{t=2}^{T}\beta_t\Big)'\right) \\
R_{j+1} &= \frac{1}{T}\left(\sum_{t=1}^{T}\alpha_t - \Big(\sum_{t=1}^{T}\delta_t\Big)\Big(\sum_{t=1}^{T}\gamma_t\Big)^{-1}\Big(\sum_{t=1}^{T}\delta_t\Big)\right).
\end{aligned}$$

2 Applying Kalman Filter to change in log output

Define:

$$y_t = log(Y_{t+1}) - log(Y_t).$$

It is given that y_t follows the following process:

$$y_t = Z\mu_t + \sigma_\epsilon \epsilon_t \tag{1}$$
$$\mu_{t+1} = B\mu_t + \sigma_\nu \nu_{t+1} \tag{2}$$

For the present case $B = [1], Z = [1], Q = \sigma_\nu^2, R = \sigma_\epsilon^2$.

2.1 Simulated State

In the gaussian case the best guess for μ_t given y^{t-1} is given by $\mathbb{E}[\mu_t|y^{t-1}]$. For the present case we set the values $\sigma_\nu = 0.15, \sigma_\epsilon = 0.10$. We plot the values for $\mathbb{E}[\mu_t|y^{t-1}]$ below:

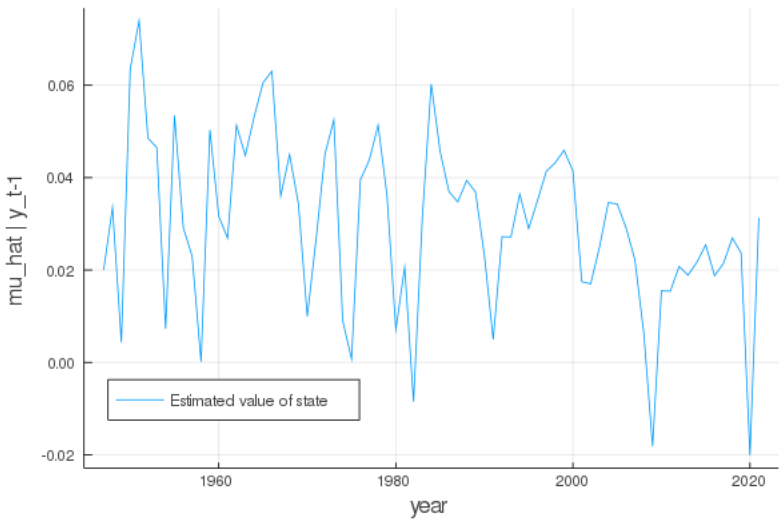

Figure 2: Estimated Value of State

2.2 Test Case

For this case we simulate the values for μ_t, y_t using the parameter values same as in the previous part. We then calculate $\mathbb{E}[\mu_t|y^{t-1}]$ and compare it with the actual values. The following figure graphs the estimates:

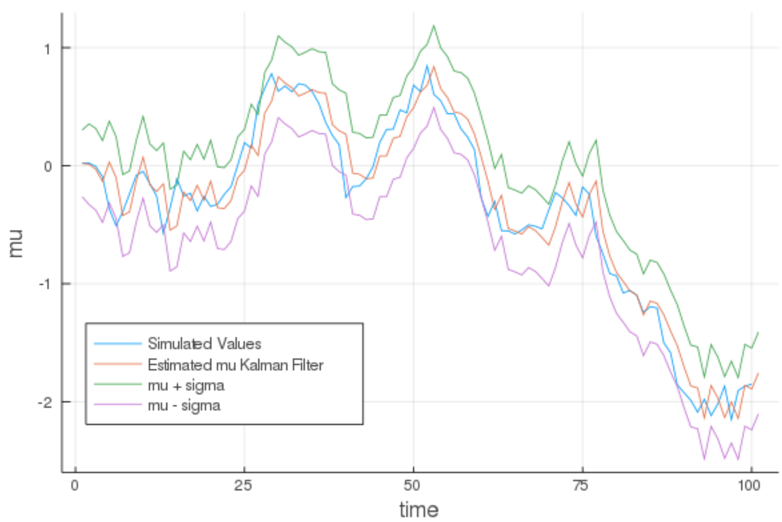

Figure 3: Estimated vs Simulated Values

3 MLE estimation

We use the following formula for MLE estimation of $sigma_\nu, \sigma_\epsilon$ based on the actual data for y_t:

$$\ln L = \sum_t \{-\frac{n}{2}\ln 2\pi - \frac{1}{2}\ln|det(F_t)| - \frac{1}{2}i_t' F_t^{-1} i_t$$

The estimates are: $\hat{\sigma}_\nu = 0.00161, \hat{\sigma}_\epsilon = 0.0224$. Using these estimates we now graph $\mathbb{E}[\hat{\mu}_{t+1}|y^{t-1}]$.

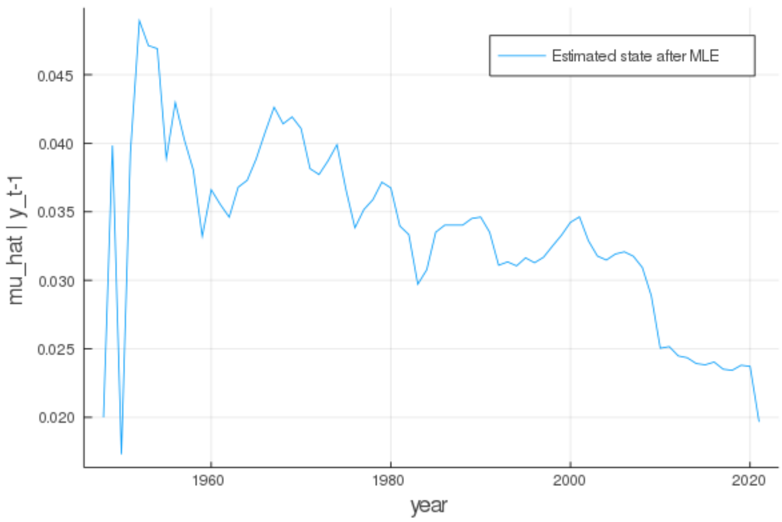

Figure 4: Estimated Value of State

4 EM Method

We follow Holmes E. E. (2012) to implement the EM algorithm. For the present case $B = [1]$ and $Z = [1]$. The EM algorithm is an iterative algorithm, where the new set of parameters maximizes the Expected value of log likelihood. The estimates from EM are: $\hat{\sigma}_\nu = 0.0196, \hat{\sigma}_\epsilon = 0.0176$.

Note: *We checked out EM algorithm for the simulated case. There both MLE estimates and EM estimates coincides. This leads us to conclude that our MLE and EM algorithm are being implemented correctly. However, the MLE and EM estimates are different when run on actual data. This may be because of low number of data points for actual growth rate data. It is also to note that the variance of growth rate is very small* (0.000547)*, which might lead to imprecise estimates.*