Lecture 8: State Space Models: Part I

Dr. Joao B. Duarte[1]

[1]Nova School of Business and Economics
University of Cambridge

Masters, Economics: Macroeconometrics
Lisbon
Spring 2017

Lecture Objectives:

- Introduction to state space models.
- ARMA and VAR casted as a state space model.
- Introduction to the Kalman Filter
- Smoothing vs Filtering
- Kalman Filter applications

Secondary Readings:

- Chapter 6, Canova
- Chapter 13, Time Series Analysis, Hamilton, James, first edition

Intro to State Space Models

- The state space models formulation is quite general.
- It encompasses all the models we have seen so far.
- However, the analysis involved are more complex and it is simpler to use the models we have seen in their previous formulations.
- However, the state space models become quite useful for different specifications. Particularly, when we are dealing with unobservable variables (state variables) and with measurement error.

Linear State Space Models

- Let the values of the state (unobserved) at time t be given by vector θ_t and y_t be a vector of observed variables at time t. The linear state space model can be represented by:

$$
\begin{aligned}
y_t &= F_t'\theta_t + \nu_t \qquad \text{(Measurement Equation)} \qquad (1) \\
\theta_t &= G_t\theta_{t-1} + w_t \qquad \text{(Transition Equation)} \qquad (2)
\end{aligned}
$$

where $\theta_0 \sim N(m_0, C_0)$, $\nu_t \sim N(0, V_t)$ and $w_t \sim N(0, W_t)$.

Examples

- Any ARMA model can be formulated as a state space model.
  Example 1: ARMA(2,1) $y_t = a_1 y_{t-1} + a_2 y_{t-1} + \varepsilon_t + b_1 \varepsilon_{t-1}$ can be written as:

(Measurement Equation)

$$y\_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ a_2 y_{t-2} + b_1 \varepsilon_{t-1} \end{bmatrix}$$

(Transition Equation)

$$\begin{bmatrix} y_t \\ a_2 y_{t-1} + b_1 \varepsilon_t \end{bmatrix} = \begin{bmatrix} a_1 & 1 \\ a_2 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ a_2 y_{t-2} + b_1 \varepsilon_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ b_1 \end{bmatrix} \varepsilon_t$$

Check!

Examples

- Any VAR model can be trivially formulated as a state space model. Example 2: VAR(1) $y_t = A_1 y_{t-1} + \varepsilon_t$ can be written as:

(Measurement Equation)

y_t = y_t

(Transition Equation)

y_t = A_1 y_t-1 + ε_t

Examples

- Any latent variable specification can also be formulated as a state space model. Example 3:

(Measurement Equation)

y_t = A_t y_t-1 + ν_{1t}

(Transition Equation)

A_t = A_t-1 + ν_{2t}

Filtering

- The State space models includes latent states (unobservable). Hence, we need to estimate the latent states in order to make predictions of the observables y_t.
- One way to estimate the latent states is called filtering.
- The idea behind filtering is to use all the information up to data t to make predictions of the state vector.

  Let D_t be all the data up to date t including the observation y_t. Then, Bayes updating gives us that:

  $$p(\theta_t|D_{t-1}, y_t) = \frac{p(y_t|\theta_t)p(\theta_t|D_{t-1})}{p(y_t)}$$

  Given the results of last lecture and the normality assumptions on the state space model we know that $\theta_t|D_t \sim N(m_t, C_t)$

Kalman Filter

- The Kalman Filter is used to compute optimal forecasts of y_t as well as recursive estimates of the state variables θ_t with time t information for state space models.
- Its possesses many useful applications:
  1. Economics:
     - Time Varying Parameters
     - Markov Switching Models
     - Unobserved Components
     - Likelihood Maximization
  2. Elsewhere:
     - Navigation
     - Signal Extraction
     - Robotics

Kalman Filter Algorithm

- The prior, likelihood and posterior distributions are given by:
  $\theta_t|D_{t-1} \sim N(a_t, R_t) \qquad y_t|\theta_t \sim N(F_t'\theta_t) \qquad \theta_t|D_t \sim N(m_t, C_t)$
- Assuming, F_t, G_t, V_t and W_t are known, the recursive algorithm of the Kalman filter, with $\theta_0 \sim N(m_0, C_0)$, is now presented:
- Step 1: Update Prior

$$\begin{aligned} a_t &= E(\theta_t|D_{t-1}) = G_t m_{t-1} && (3) \\ R_t &= var(\theta_t|D_{t-1}) = G_t C_{t-1} G_t' + W_t && (4) \end{aligned}$$

Kalman Filter Algorithm

- Step 2: Forecast y_t and mean square of the forecast error (with $t-1$ info)

$$\begin{aligned} f_t &= E(y_t|D_{t-1}) = F_t' a_t && (5) \\ Q_t &= var(y_t|D_{t-1}) = F_t' R_t F_t + V_t && (6) \end{aligned}$$

- Step 3: Calculate the prediction error and the Kalman gain

$$\begin{aligned} e_t &= y_t - f_t && (7) \\ K_t &= R_t F_t Q_t^{-1} && (8) \end{aligned}$$

Kalman Filter Algorithm

- Step 4: Update state estimate: (with t information)

$$
\begin{aligned}
m_t &= a_t + K_t e_t && (9)\\
C_t &= R_t - K_t Q_t K_t' && (10)
\end{aligned}
$$

- Step 5: Repeat previous steps until $t = T$
- The posterior mean of the state is a weighted sum of the prior mean and the forecast error.
- Also, notice that the variance of the posterior distribution, C_t, is less than the variance of the prior distribution R_t.

Smoothing

- Another way to estimate the state vector is to use the entire sample information instead of just up to time t.
- The strategy is then to start in the lat period observation and update the state θ_t backwards.

$$p(\theta_{t-1}|D_t) = \int p(\theta_{t-1}|\theta_t, D_t)p(\theta_t|D_t)d\theta_t$$

$$\theta_{t-1}|D_t \sim N(a_t(-1), R_t(-1))$$

where

$$\begin{aligned} a_t(-1) &= m_{t-1} + B_{t-1}(m_t - a_t) \\ R_t(-1) &= C_{t-1} - B_{t-1}(R_t - C_t)B'_{t-1} \\ B_{t-1} &= C_{t-1}G'_tR_t^{-1} \end{aligned}$$

Smoothing vs Filtering

- Smoothing is particularly useful when we are interested in the value of the unobserved variables for a specific sample period.
- It is not useful to make predictions of observables.
- Important: Cannot be used to estimate likelihood and model parameters. For that, we need to use filtering.

Prediction

- Theorem: If initial conditions (priors) and innovations are normal, Kalman filter is the best predictor (linear and nonlinear) of y_t. Else, it gives the best linear predictor.
- With the estimated state at date t using the Kalman Filter we can make predictions of future values of observables.
- The one-step ahead forecast is given by:

$$E[y_{t+1}|D_t] = E[F'_{t+1}\theta_{t+1} + \nu_{t+1}|D_t] = F'_{t+1}a_{t+1} = f_{t+1}$$

$$var[y_{t+1}|D_t] = var[F'_{t+1}\theta_{t+1} + \nu_{t+1}|D_t] =$$

F'_t+1var$[\theta_{t+1}|D_t]F_{t+1} + V_{t+1} = F'_{t+1}R_{t+1}F_{t+1} + V_{t+1} = Q_{t+1}$

Maximum Likelihood Estimation

- Let y_t be an $m \times 1$ and let ϕ represent a vector of coefficients to be estimated. The likelihood $\mathcal{L}(y_1, ..., y_t, \phi)$ can be written as a decomposition of the prediction error:

$$\begin{aligned}\mathcal{L}(y_1, ..., y_t, \phi) = & -\frac{Tm}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{t} \ln |\Sigma_{t|t-1}| \\ & - \frac{1}{2} \sum_{t=1}^{T} (y_t - y_{t|t-1}) \Sigma_{t|t-1}^{-1} (y_t - y_{t|t-1})\end{aligned} \tag{11}$$

  where $e_t = y_t - y_{t|t-1} \sim N(0, \Sigma_{t|t-1})$ and $y_1 \sim N(\bar{y}_1, \Sigma_1)$; $e_1 = y_1 - \bar{y}_1$

- Given ϕ the Kalman filter can be used to compute $e_t, \Sigma_{t|t-1}$ for all t. Then we can estimate ϕ by maximizing (11). Do this until convergence. This process is called the EM algorithm.

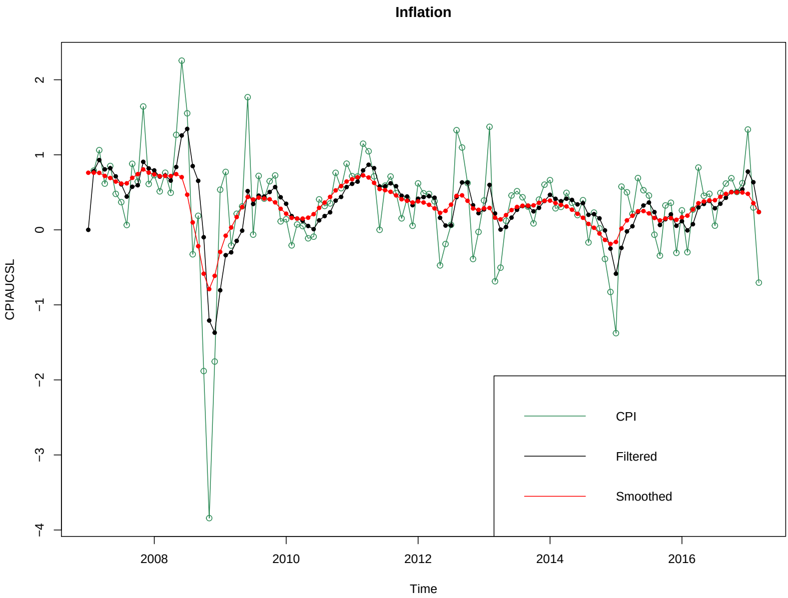
Applications

- Random Walk plus noise model

- As an example, consider the monthly inflation of CPI.
- In practice, we do not observe true inflation. Moreover, even the inflation CPI data has some measurement error.
- We can model inflation using the following state space model:

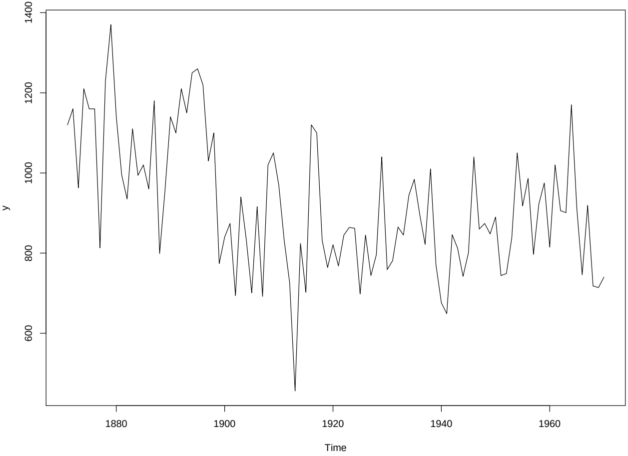
$$
\begin{aligned}
CPI_t &= \pi_t + \nu_t \\
\pi_t &= \pi_{t-1} + w_t
\end{aligned}
$$

Applications



Applications

- Now consider the Nile flow at Aswan data from 1871 to 1970.



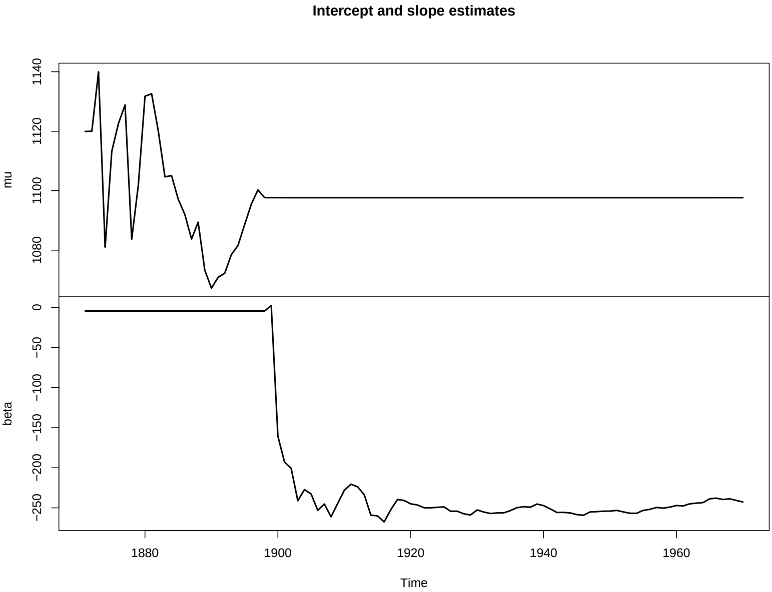
Applications

- We could try to model the moving average flow of the river using a state space model with time varying coefficients to capture the structural change that seemed to happen around 1900.
- In order to do that, we will introduce a time dummy variable x_t that will take value 0 just before 1900 and 1 thereafter.

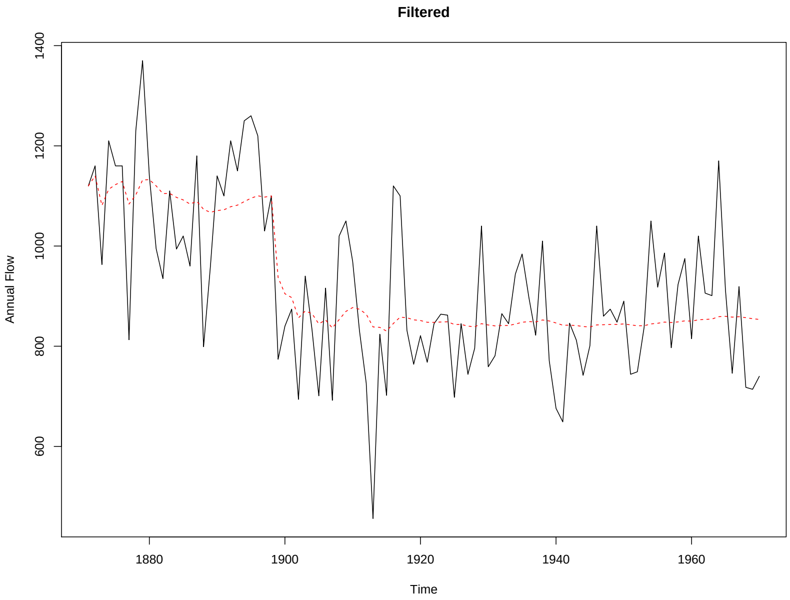
$$
\begin{aligned}
y_t &= \mu_t + \lambda_t x_t + \nu_t \\
\mu_t &= \mu_{t-1} + w_{1t} \\
\lambda_t &= \lambda_{t-1} + w_{2t}
\end{aligned}
$$

- In this example, F_t varies over time and our state vector is $\theta_t = (\mu_t, \lambda_t)^T$

Applications



Applications



Questions to think about

- Although state space models are fairly general as presented in this lecture, what extensions could make them even more general?
- What is a latent variable? How can we optimally estimate a latent variable?
- What is the basic idea behind the Kalman Filter?
- Why can't we use smoothing to estimate likelihood?