BANK OF ENGLAND

Centre for Central Banking Studies

State-space models and the Kalman filter

Date

Nairobi, April 26, 2016

Title

Pawel Zabczyk

pawel.zabczyk@bankofengland.co.uk

State space models are useful/crucial for applied research

Can be used to

- measure how a relationship between variables changes over time
- decompose a time-series into a cycle and trend component
- determine if a common component is driving a group of time series
- calculate the likelihood function for many models (including DSGE's)

State space models are useful/crucial for applied research

Can be used to

- measure how a relationship between variables changes over time
- decompose a time-series into a cycle and trend component
- determine if a common component is driving a group of time series
- calculate the likelihood function for many models (including DSGE's)

State space models are useful/crucial for applied research

Can be used to

- measure how a relationship between variables changes over time
- decompose a time-series into a cycle and trend component
- determine if a common component is driving a group of time series
- calculate the likelihood function for many models (including DSGE's)

State space models are useful/crucial for applied research

Can be used to

- measure how a relationship between variables changes over time
- decompose a time-series into a cycle and trend component
- determine if a common component is driving a group of time series
- calculate the likelihood function for many models (including DSGE's)

State space models

Observation equation

- $\underset{\text{Data}}{y_t} = \underset{\text{Coefficient/Data}}{x_t} \overset{\text{State}}{\beta_t} + e_t$

Transition equation

- $\beta_t = \mu_t + F\beta_{t-1} + v_t$

We assume

- $e_t \sim i.i.d.N(0, R)$
- $v_t \sim i.i.d.N(0, Q)$
- $\forall t, s \; E(e_t v_s') = 0$

where R and Q are variance-covariance matrices

State space models

Observation equation

- $\underset{\text{Data}}{y_t} = \underset{\text{Coefficient/Data}}{x_t} \; \overset{\text{State}}{\beta_t} + e_t$

Transition equation

- $\beta_t = \mu_t + F\beta_{t-1} + v_t$

We assume

- $e_t \sim i.i.d.N(0, R)$
- $v_t \sim i.i.d.N(0, Q)$
- $\forall t, s \; E(e_t v_s') = 0$

where R and Q are variance-covariance matrices

State space models

Observation equation

- $\underset{\text{Data}}{y_t} = \underset{\text{Coefficient/Data}}{x_t} \overset{\text{State}}{\beta_t} + e_t$

Transition equation

- $\beta_t = \mu_t + F\beta_{t-1} + v_t$

We assume

- $e_t \sim i.i.d.N(0, R)$
- $v_t \sim i.i.d.N(0, Q)$
- $\forall t, s \; E(e_t v_s') = 0$

where R and Q are variance-covariance matrices

State space models

Observation equation

- Data y_t = Coefficient/Data x_t State $\beta_t + e_t$

Transition equation

- $\beta_t = \mu_t + F\beta_{t-1} + v_t$

We assume

- $e_t \sim i.i.d.N(0, R)$
- $v_t \sim i.i.d.N(0, Q)$
- $\forall t, s\ E(e_t v_s') = 0$

where R and Q are variance-covariance matrices

State space models

Observation equation

- $\underset{\text{Data}}{y_t} = \underset{\text{Coefficient/Data}}{x_t} \overset{\text{State}}{\beta_t} + e_t$

Transition equation

- $\beta_t = \mu_t + F\beta_{t-1} + v_t$

We assume

- $e_t \sim i.i.d.N(0, R)$
- $v_t \sim i.i.d.N(0, Q)$
- $\forall t, s \; E(e_t v_s') = 0$

where R and Q are variance-covariance matrices

State space models

Observation equation

- $\underset{\text{Data}}{y_t} = \underset{\text{Coefficient/Data}}{x_t} \overset{\text{State}}{\beta_t} + e_t$

Transition equation

- $\beta_t = \mu_t + F\beta_{t-1} + v_t$

We assume

- $e_t \sim i.i.d.N(0, R)$
- $v_t \sim i.i.d.N(0, Q)$
- $\forall t, s \; E(e_t v_s') = 0$

where R and Q are variance-covariance matrices

Examples: A Time-varying parameter regression

$Y_t = c_t + Z_t B_t + e_t$, c_t and B_t follow a random walk

- How to write this in State Space form?

Examples: A Time-varying parameter regression

$Y_t = c_t + Z_t B_t + e_t$, c_t and B_t follow a random walk

- How to write this in State Space form?

$$Y_t = \overset{x_t}{\begin{pmatrix} 1 & Z_t \end{pmatrix}} \overset{\beta_t}{\begin{pmatrix} c_t \\ B_t \end{pmatrix}} + e_t, VAR(e_t) = R$$

$$\overset{\beta_t}{\begin{pmatrix} c_t \\ B_t \end{pmatrix}} = \overset{\beta_{t-1}}{\begin{pmatrix} c_{t-1} \\ B_{t-1} \end{pmatrix}} + \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}, VAR(v_t) = Q$$

Note: $F = I$ and $\mu = 0$

Examples: A Trend Cycle Model

Decomposing GDP into trend and cycle

- Let $Y_t = C_t + \tau_t$ where
  - the cycle C_t follows an AR (2) process
    $C_t = c + \rho_1 C_{t-1} + \rho_2 C_{t-2} + v_{1t}$
  - the trend follows a random walk
    $\tau_t = \tau_{t-1} + v_{2t}$
- How to write this in state space form?

Examples: A Trend Cycle Model

Decomposing GDP into trend and cycle

- Let $Y_t = C_t + \tau_t$ where
  - the cycle C_t follows an AR (2) process $C_t = c + \rho_1 C_{t-1} + \rho_2 C_{t-2} + v_{1t}$
  - the trend follows a random walk $\tau_t = \tau_{t-1} + v_{2t}$
- How to write this in state space form?

Examples: A Trend Cycle Model

Decomposing GDP into trend and cycle

- Let $Y_t = C_t + \tau_t$ where
  - the cycle C_t follows an AR (2) process
    $C_t = c + \rho_1 C_{t-1} + \rho_2 C_{t-2} + v_{1t}$
  - the trend follows a random walk
    $\tau_t = \tau_{t-1} + v_{2t}$

- How to write this in state space form?

Examples: A Trend Cycle Model

Decomposing GDP into trend and cycle

- Let $Y_t = C_t + \tau_t$ where
  - the cycle C_t follows an AR (2) process
    $C_t = c + \rho_1 C_{t-1} + \rho_2 C_{t-2} + v_{1t}$
  - the trend follows a random walk
    $\tau_t = \tau_{t-1} + v_{2t}$
- How to write this in state space form?

Examples: A Trend Cycle Model

Decomposing GDP into trend and cycle

- Let $Y_t = C_t + \tau_t$ where
  - the cycle C_t follows an AR (2) process
    $C_t = c + \rho_1 C_{t-1} + \rho_2 C_{t-2} + v_{1t}$
  - the trend follows a random walk
    $\tau_t = \tau_{t-1} + v_{2t}$
- How to write this in state space form?

$$Y_t = \overset{x}{\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}} \overset{\beta_t}{\begin{pmatrix} C_t \\ \tau_t \\ C_{t-1} \end{pmatrix}}; \quad R \equiv 0$$

$$\overset{\beta_t}{\begin{pmatrix} C_t \\ \tau_t \\ C_{t-1} \end{pmatrix}} = \overset{\mu}{\begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix}} + \overset{F}{\begin{pmatrix} \rho_1 & 0 & \rho_2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}} \overset{\beta_{t-1}}{\begin{pmatrix} C_{t-1} \\ \tau_{t-1} \\ C_{t-2} \end{pmatrix}} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ 0 \end{pmatrix}$$

Examples: A Dynamic factor model

A panel of series Y_{it} has a common component

$$Y_{it} = B_i F_t + e_{it}$$

where $F_t = c + \rho_1 F_{t-1} + \rho_2 F_{t-2} + v_t$

- How to write this in state space form?

Examples: A Dynamic factor model

A panel of series Y_{it} has a common component

$$Y_{it} = B_i F_t + e_{it}$$

where $F_t = c + \rho_1 F_{t-1} + \rho_2 F_{t-2} + v_t$

- How to write this in state space form?

$$\overset{y_t}{\begin{pmatrix} Y_{1t} \\ Y_{2t} \\ . \\ Y_{Nt} \end{pmatrix}} = \overset{x}{\begin{pmatrix} B_1 & 0 \\ B_1 & 0 \\ . & . \\ B_N & 0 \end{pmatrix}} \overset{\beta_t}{\begin{pmatrix} F_t \\ F_{t-1} \end{pmatrix}} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ . \\ e_{Nt} \end{pmatrix}$$

$$\overset{\beta_t}{\begin{pmatrix} F_t \\ F_{t-1} \end{pmatrix}} = \overset{\mu}{\begin{pmatrix} c \\ 0 \end{pmatrix}} + \overset{F}{\begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix}} \overset{\beta_{t-1}}{\begin{pmatrix} F_{t-1} \\ F_{t-2} \end{pmatrix}} + \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}$$

Examples: Conditional Forecasting

- Imagine you've just estimated the following VAR

$$\underset{\beta_t}{\begin{pmatrix} GDP_t \\ i_t \end{pmatrix}} = \underset{\mu}{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}} + \underset{F}{\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}} \underset{\beta_{t-1}}{\begin{pmatrix} GDP_{t-1} \\ i_{t-1} \end{pmatrix}} + \underset{v_t}{\begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}}$$

where i_t denotes the nominal interest rate. You could use the state-space toolkit to forecast GDP conditional on any path of interest rates $\bar{i}_{cf,t+1}, , \ldots, \bar{i}_{cf,t+n}$

Examples: Conditional Forecasting

- Imagine you've just estimated the following VAR

$$\underbrace{\begin{pmatrix} GDP_t \\ i_t \end{pmatrix}}_{\beta_t} = \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{\mu} + \underbrace{\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} GDP_{t-1} \\ i_{t-1} \end{pmatrix}}_{\beta_{t-1}} + \underbrace{\begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}}_{v_t}$$

where i_t denotes the nominal interest rate. You could use the state-space toolkit to forecast GDP conditional on any path of interest rates $\bar{i}_{cf,t+1},, \ldots, \bar{i}_{cf,t+n}$

- To do that take μ and F from the VAR and set-up

$$\underbrace{\left(\bar{i}_{cf,t+s}\right)}_{y_s} = \underbrace{\begin{pmatrix} 0 & 1 \end{pmatrix}}_{x} \underbrace{\begin{pmatrix} GDP_{t+s} \\ i_{t+s} \end{pmatrix}}_{\beta_s}$$

$$\underbrace{\begin{pmatrix} GDP_{t+s} \\ i_{t+s} \end{pmatrix}}_{\beta_s} = \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{\mu} + \underbrace{\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} GDP_{t+s-1} \\ i_{t+s-1} \end{pmatrix}}_{\beta_{s-1}} + \underbrace{\begin{pmatrix} v_{1t+s} \\ v_{2t+s} \end{pmatrix}}_{v_s}$$

Examples: Interpolation of data

- Suppose we have quarterly *GDP* and we want to estimate a monthly series $\hat{Y}_t$ using information in monthly data Z_t. We can treat monthly GDP numbers as unobserved states!
- Specifically, we could assume

$$\hat{Y}_t = \theta \hat{Y}_{t-1} + DZ_t + u_t, u_t = \rho u_{t-1} + g_t, VAR(g) = \sigma^2$$

and define

$$y \text{ as } \begin{pmatrix} 0 \\ 0 \\ GDP_3 \\ 0 \\ 0 \\ GDP_6 \\ . \end{pmatrix} \text{ with } x \text{ given by } \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ . & . & . & . \end{pmatrix}$$

where the unobserved state $\beta_t = \begin{pmatrix} \hat{Y}_t, & \hat{Y}_{t-1}, & \hat{Y}_{t-2}, & u_t \end{pmatrix}'$.

Examples: Interpolation of data

- Suppose we have quarterly *GDP* and we want to estimate a monthly series $\hat{Y}_t$ using information in monthly data Z_t. We can treat monthly GDP numbers as unobserved states!
- Specifically, we could assume

$$\hat{Y}_t = \theta \hat{Y}_{t-1} + DZ_t + u_t, u_t = \rho u_{t-1} + g_t, VAR(g) = \sigma^2$$

and define

$$y \text{ as } \begin{pmatrix} 0 \\ 0 \\ GDP_3 \\ 0 \\ 0 \\ GDP_6 \\ . \end{pmatrix} \text{ with } x \text{ given by } \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ . & . & . & . \end{pmatrix}$$

where the unobserved state $\beta_t = \begin{pmatrix} \hat{Y}_t, & \hat{Y}_{t-1}, & \hat{Y}_{t-2}, & u_t \end{pmatrix}'$.

Examples: Interpolation of data

The transition equation would then be given by

$$\underbrace{\begin{pmatrix} \hat{Y}_t \\ \hat{Y}_{t-1} \\ \hat{Y}_{t-2} \\ u_t \end{pmatrix}}_{\beta_t} = \underbrace{\begin{pmatrix} DZ_t \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\mu_t} + \underbrace{\begin{pmatrix} \theta & 0 & 0 & \rho \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \rho \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} \hat{Y}_{t-1} \\ \hat{Y}_{t-2} \\ \hat{Y}_{t-3} \\ u_{t-1} \end{pmatrix}}_{\beta_{t-1}} + \begin{pmatrix} g_t \\ 0 \\ 0 \\ g_t \end{pmatrix}$$

with $Q = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}$

Estimation Aims

- Observation equation $y_t = \mathbf{x}_t \beta_t + e_t$, $VAR(e) = \mathbf{R}$
- Transition equation $\beta_t = \boldsymbol{\mu}_t + \mathbf{F}\beta_{t-1} + v_t$, $VAR(v) = \mathbf{Q}$
- Parameters of the state space **and** the state variables are unknown

Estimation Aims

- Observation equation $y_t = \mathbf{x}_t \beta_t + e_t, VAR(e) = \mathbf{R}$
- Transition equation $\beta_t = \boldsymbol{\mu}_t + \mathbf{F}\beta_{t-1} + v_t, VAR(v) = \mathbf{Q}$
- Parameters of the state space **and** the state variables are unknown

Estimation Aims

- Observation equation $y_t = \mathbf{x}_t\beta_t + e_t, VAR(e) = \mathbf{R}$
- Transition equation $\beta_t = \boldsymbol{\mu}_t + \mathbf{F}\beta_{t-1} + v_t, VAR(v) = \mathbf{Q}$
- Parameters of the state space **and** the state variables are unknown

Estimating the states via the Kalman filter

Consider the state space formulation

$$
\begin{aligned}
y_t &= x_t\beta_t + e_t, e_t \sim N(0, R) \\
\beta_t &= \mu + F\beta_{t-1} + v_t, v_t \sim N(0, Q)
\end{aligned}
$$

Notation: $\beta_{t\backslash t}$ inference on β_t given information up to time **t**
Notation: $P_{t\backslash t}$ covariance of β_t given information up to time **t**

- Assume for the moment that R, Q, μ, F are known

The Kalman filter (Overview)

Starting Values (time 0)

$$\beta_{0\backslash 0}, P_{0\backslash 0}$$

Predict States (time 1...onwards)

$$\beta_{t\backslash t-1} = \mu + F\beta_{t-1\backslash t-1}$$
$$P_{t\backslash t-1} = FP_{t-1\backslash t-1}F' + Q$$

Calculate Prediction Error

$$\eta_{t\backslash t-1} = y_t - x_t\beta_{t\backslash t-1}$$
$$f_{t\backslash t-1} = x_t P_{t\backslash t-1} x_t' + R$$

Update States

$$K = P_{t\backslash t-1} x_t' f_{t\backslash t-1}^{-1}$$
$$\beta_{t\backslash t} = \beta_{t\backslash t-1} + K\eta_{t\backslash t-1}$$
$$P_{t\backslash t} = P_{t\backslash t-1} - Kx_t P_{t\backslash t-1}$$

repeat for $t = 1 \ldots T$

The Prediction equations of the Kalman filter

- First equation $\beta_{t\backslash t-1} = \mu + F\beta_{t-1\backslash t-1}$. This is simply the expected value of the transition equation $E\left(\mu + F\beta_{t-1} + v_t\right)$
- Second equation $P_{t\backslash t-1} = FP_{t-1\backslash t-1}F' + Q$. Can be derived as $VAR\left(\mu + F\beta_{t-1} + v_t\right)$
- Third equation of the Kalman filter just compares the outcome to the prediction: $\eta_{t\backslash t-1} = y_t - x_t\beta_{t\backslash t-1}$
- Fourth equation of the Kalman filter: $f_{t\backslash t-1} = x_t P_{t\backslash t-1} x_t' + R$ Can be derived as

$$E([x_t\beta_t + v_t] - x_t\beta_{t\backslash t-1})^2 = E(x_t\left(\beta_t - \beta_{t\backslash t-1}\right)^2 x_t' + v_t v_t')$$

The Prediction equations of the Kalman filter

- First equation $\beta_{t\backslash t-1} = \mu + F\beta_{t-1\backslash t-1}$. This is simply the expected value of the transition equation $E\left(\mu + F\beta_{t-1} + v_t\right)$
- Second equation $P_{t\backslash t-1} = FP_{t-1\backslash t-1}F' + Q$. Can be derived as $VAR\left(\mu + F\beta_{t-1} + v_t\right)$
- Third equation of the Kalman filter just compares the outcome to the prediction: $\eta_{t\backslash t-1} = y_t - x_t\beta_{t\backslash t-1}$
- Fourth equation of the Kalman filter: $f_{t\backslash t-1} = x_t P_{t\backslash t-1} x_t' + R$ Can be derived as

  $E([x_t\beta_t + v_t] - x_t\beta_{t\backslash t-1})^2 = E(x_t\left(\beta_t - \beta_{t\backslash t-1}\right)^2 x_t' + v_t v_t')$

The Prediction equations of the Kalman filter

- First equation $\beta_{t\backslash t-1} = \mu + F\beta_{t-1\backslash t-1}$. This is simply the expected value of the transition equation $E\left(\mu + F\beta_{t-1} + v_t\right)$
- Second equation $P_{t\backslash t-1} = FP_{t-1\backslash t-1}F' + Q$. Can be derived as $VAR\left(\mu + F\beta_{t-1} + v_t\right)$
- Third equation of the Kalman filter just compares the outcome to the prediction: $\eta_{t\backslash t-1} = y_t - x_t\beta_{t\backslash t-1}$
- Fourth equation of the Kalman filter: $f_{t\backslash t-1} = x_t P_{t\backslash t-1} x_t' + R$ Can be derived as $E([x_t\beta_t + v_t] - x_t\beta_{t\backslash t-1})^2 = E(x_t\left(\beta_t - \beta_{t\backslash t-1}\right)^2 x_t' + v_t v_t')$

The Prediction equations of the Kalman filter

- First equation $\beta_{t\backslash t-1} = \mu + F\beta_{t-1\backslash t-1}$. This is simply the expected value of the transition equation $E\left(\mu + F\beta_{t-1} + v_t\right)$
- Second equation $P_{t\backslash t-1} = FP_{t-1\backslash t-1}F' + Q$. Can be derived as $VAR\left(\mu + F\beta_{t-1} + v_t\right)$
- Third equation of the Kalman filter just compares the outcome to the prediction: $\eta_{t\backslash t-1} = y_t - x_t\beta_{t\backslash t-1}$
- Fourth equation of the Kalman filter: $f_{t\backslash t-1} = x_t P_{t\backslash t-1} x_t' + R$ Can be derived as $E([x_t\beta_t + v_t] - x_t\beta_{t\backslash t-1})^2 = E(x_t\left(\beta_t - \beta_{t\backslash t-1}\right)^2 x_t' + v_t v_t')$

The Updating equations of the Kalman filter

- Fifth equation: $\beta_{t\backslash t} = \beta_{t\backslash t-1} + K\eta_{t\backslash t-1}$ where $K = P_{t\backslash t-1} x_t' f_{t\backslash t-1}^{-1}$
- Updates old estimate $\beta_{t\backslash t-1}$ with information contained in $\eta_{t\backslash t-1}$
- The Kalman gain K is weight given to new information
- The prediction error contains information that is new relative to the past data

The Updating equations of the Kalman filter

- Fifth equation: $\beta_{t\backslash t} = \beta_{t\backslash t-1} + K\eta_{t\backslash t-1}$ where $K = P_{t\backslash t-1} x_t' f_{t\backslash t-1}^{-1}$
- Updates old estimate $\beta_{t\backslash t-1}$ with information contained in $\eta_{t\backslash t-1}$
- The Kalman gain K is weight given to new information
- The prediction error contains information that is new relative to the past data

The Updating equations of the Kalman filter

- Fifth equation: $\beta_{t\backslash t} = \beta_{t\backslash t-1} + K\eta_{t\backslash t-1}$ where $K = P_{t\backslash t-1} x_t' f_{t\backslash t-1}^{-1}$
- Updates old estimate $\beta_{t\backslash t-1}$ with information contained in $\eta_{t\backslash t-1}$
- The Kalman gain K is weight given to new information
- The prediction error contains information that is new relative to the past data

The Updating equations of the Kalman filter

- Fifth equation: $\beta_{t\backslash t} = \beta_{t\backslash t-1} + K\eta_{t\backslash t-1}$ where $K = P_{t\backslash t-1}x_t' f_{t\backslash t-1}^{-1}$
- Updates old estimate $\beta_{t\backslash t-1}$ with information contained in $\eta_{t\backslash t-1}$
- The Kalman gain K is weight given to new information
- The prediction error contains information that is new relative to the past data

The Updating equations of the Kalman filter

- The Kalman gain in a univariate setting is given by

$$K = \frac{1}{x_t} \frac{P_{t\backslash t-1} x_t^2 \text{ [due to uncertainty about state]}}{P_{t\backslash t-1} x_t^2 + R \text{ [due to uncertainty about shock]}}$$

- Note that K increases as $P_{t\backslash t-1} x_t^2$ rises and more weight is placed on new information in $\eta_{t\backslash t-1}$
- K falls as R increases and the shock is less informative

The Updating equations of the Kalman filter

- The Kalman gain in a univariate setting is given by

$$K = \frac{1}{x_t} \frac{P_{t \backslash t-1} x_t^2 \text{ [due to uncertainty about state]}}{P_{t \backslash t-1} x_t^2 + R \text{ [due to uncertainty about shock]}}$$

- Note that K increases as $P_{t \backslash t-1} x_t^2$ rises and more weight is placed on new information in $\eta_{t \backslash t-1}$
- K falls as R increases and the shock is less informative

The Updating equations of the Kalman filter

- The Kalman gain in a univariate setting is given by

$$K = \frac{1}{x_t} \frac{P_{t\backslash t-1} x_t^2 \text{ [due to uncertainty about state]}}{P_{t\backslash t-1} x_t^2 + R \text{ [due to uncertainty about shock]}}$$

- Note that K increases as $P_{t\backslash t-1} x_t^2$ rises and more weight is placed on new information in $\eta_{t\backslash t-1}$
- K falls as R increases and the shock is less informative

The initial values

- For a stationary transition equation

$$\beta_{0\backslash 0} = (I - F)^{-1} \mu$$
$$P_{0\backslash 0} = (I - F \otimes F)^{-1} vec(Q)$$

- For the non-stationary case

$$\beta_{0\backslash 0} = \text{arbitrary}$$
$$P_{0\backslash 0} = \text{large number}$$

The initial values

- For a stationary transition equation

$$\beta_{0\backslash 0} = (I - F)^{-1} \mu$$
$$P_{0\backslash 0} = (I - F \otimes F)^{-1} vec(Q)$$

- For the non-stationary case

$$\beta_{0\backslash 0} = \text{arbitrary}$$
$$P_{0\backslash 0} = \text{large number}$$

Kalman Smoothing

- The Kalman filter provides estimates of the state given current information
- One may be interested in estimating the state vector at date t based on information contained in the entire data set
- Kalman smoothing updates $\beta_{t \backslash t}$ for this information, it calculates $E\left(\beta_t \backslash \beta_{t+1}, y_t\right)$ for each t

Kalman Smoothing

- The Kalman filter provides estimates of the state given current information
- One may be interested in estimating the state vector at date t based on information contained in the entire data set
- Kalman smoothing updates $\beta_{t\backslash t}$ for this information, it calculates $E\left(\beta_t \backslash \beta_{t+1}, y_t\right)$ for each t

Kalman Smoothing

- The Kalman filter provides estimates of the state given current information
- One may be interested in estimating the state vector at date t based on information contained in the entire data set
- Kalman smoothing updates $\beta_{t\backslash t}$ for this information, it calculates $E\left(\beta_t \backslash \beta_{t+1}, y_t\right)$ for each t

Kalman Smoothing

- First run the Kalman filter and save $\beta_{t+1\backslash t}$, $\beta_{t\backslash t}$, $P_{t+1/t}$, $P_{t/t}$

Starting values (time T)

$$\beta_{T/T}, P_{T/T}$$

For T=t-1,...1

Update states

$$\beta_{t\backslash T} = \beta_{t\backslash t} + P_{t\backslash t} F' P^{-1}_{t+1\backslash t}(\beta_{t+1\backslash T} - F\beta_{t\backslash t} - \mu)$$

$$P_{t\backslash T} = P_{t\backslash t} + P_{t\backslash t} F' P^{-1}_{t+1\backslash t}(P_{t+1\backslash T} - P_{t+1\backslash t}) P^{-1'}_{t+1\backslash t} F P'_{t\backslash t}$$

Kalman Smoothing

Starting values (time T) $\beta_{T/T}, P_{T/T}$

↓

For T=t-1,...1

Kalman filtered state — gain — Prediction error

Update states

$$\beta_{t\backslash T} = \beta_{t\backslash t} + P_{t\backslash t}F'P^{-1}_{t+1\backslash t}(\beta_{t+1\backslash T} - F\beta_{t\backslash t} - \mu)$$
$$P_{t\backslash T} = P_{t\backslash t} + P_{t\backslash t}F'P^{-1}_{t+1\backslash t}(P_{t+1\backslash T} - P_{t+1\backslash t})P^{-1\prime}_{t+1\backslash t}FP'_{t\backslash t}$$

- Knowledge of y_{t+j}, x_{t+j} redundant as this is already incorporated in β_{t+1}

Example of filtering and smoothing

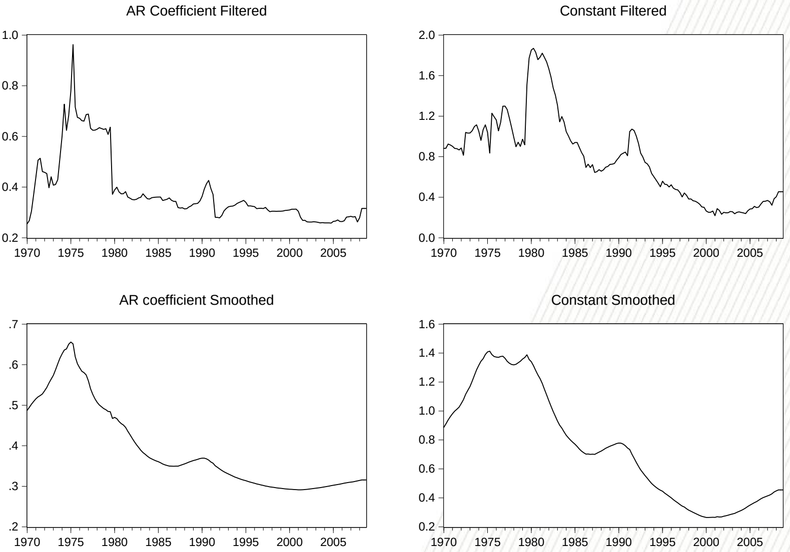
Consider the following Time-varying regression for UK inflation

$$\pi_t = c_t + b_t \pi_{t-1} + \varepsilon_t, VAR(\varepsilon_t) = R = 0.6$$

$$\underbrace{\begin{pmatrix} c_t \\ b_t \end{pmatrix}}_{\beta_t} = \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\mu} + \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} c_{t-1} \\ b_{t-1} \end{pmatrix}}_{\beta_{t-1}} + \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix},$$

$$VAR \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} = Q = \begin{pmatrix} 0.006 & 0 \\ 0 & 0.001 \end{pmatrix}$$

Example of filtering and smoothing



Maximum Likelihood estimation

- R, Q, μ and F are generally not known but need to be estimated
- The Kalman filter provides us with an estimate of the likelihood function of the model
- Recall

$$
\begin{aligned}
y_t &= \mathbf{x}\beta_t + e_t,\ VAR(e) = \mathbf{R} \\
\beta_t &= \mu + \mathbf{F}\beta_{t-1} + v_t,\ VAR(v) = \mathbf{Q}
\end{aligned}
$$

- Assuming normal error terms $\rightarrow$

$$y_t \backslash x_t \sim N\left(\mathbf{x}\beta_t, xP_{t\backslash t-1}x' + R\right)$$

- The likelihood function for each observation is

$$(2\pi)^{-n/2}\left|xP_{t\backslash t-1}x' + R\right|^{-0.5}$$

$$\exp\left(-0.5\left(y_t - x\beta_t\right)'\left(xP_{t\backslash t-1}x' + R\right)^{-1}\left(y_t - x\beta_t\right)\right)$$

$$or\ \ (2\pi)^{-n/2}\left|f_{t\backslash t-1}\right|^{-0.5}\exp\left(-0.5\eta'_{t\backslash t-1}\left(f_{t\backslash t-1}\right)^{-1}\eta_{t\backslash t-1}\right)$$

Maximum Likelihood estimation

- R, Q, μ and F are generally not known but need to be estimated
- The Kalman filter provides us with an estimate of the likelihood function of the model
- Recall
$$\begin{aligned} y_t &= \mathbf{x}\beta_t + e_t, VAR(e) = \mathbf{R} \\ \beta_t &= \mu + \mathbf{F}\beta_{t-1} + v_t, VAR(v) = \mathbf{Q} \end{aligned}$$
- Assuming normal error terms $\rightarrow$
$$y_t \backslash x_t \sim N\left(\mathbf{x}\beta_t, xP_{t\backslash t-1}x' + R\right)$$
- The likelihood function for each observation is

$$(2\pi)^{-n/2}\left|xP_{t\backslash t-1}x' + R\right|^{-0.5}$$
$$\exp\left(-0.5\left(y_t - x\beta_t\right)'\left(xP_{t\backslash t-1}x' + R\right)^{-1}\left(y_t - x\beta_t\right)\right)$$

or $(2\pi)^{-n/2}\left|f_{t\backslash t-1}\right|^{-0.5}\exp\left(-0.5\eta'_{t\backslash t-1}\left(f_{t\backslash t-1}\right)^{-1}\eta_{t\backslash t-1}\right)$

Maximum Likelihood estimation

- R, Q, μ and F are generally not known but need to be estimated
- The Kalman filter provides us with an estimate of the likelihood function of the model
- Recall

$$\begin{aligned} y_t &= \mathbf{x}\beta_t + e_t, VAR(e) = \mathbf{R} \\ \beta_t &= \boldsymbol{\mu} + \mathbf{F}\beta_{t-1} + v_t, VAR(v) = \mathbf{Q} \end{aligned}$$

- Assuming normal error terms $\rightarrow$

$$y_t \backslash x_t \sim N\left(\mathbf{x}\beta_t, xP_{t\backslash t-1}x' + R\right)$$

- The likelihood function for each observation is

$$(2\pi)^{-n/2}\left|xP_{t\backslash t-1}x' + R\right|^{-0.5} \exp\left(-0.5\left(y_t - x\beta_t\right)'\left(xP_{t\backslash t-1}x' + R\right)^{-1}\left(y_t - x\beta_t\right)\right)$$

$$or \;\; (2\pi)^{-n/2}\left|f_{t\backslash t-1}\right|^{-0.5} \exp\left(-0.5\eta'_{t\backslash t-1}\left(f_{t\backslash t-1}\right)^{-1}\eta_{t\backslash t-1}\right)$$

Maximum Likelihood estimation

- R, Q, μ and F are generally not known but need to be estimated
- The Kalman filter provides us with an estimate of the likelihood function of the model
- Recall

$$\begin{aligned} y_t &= \mathbf{x}\beta_t + e_t, VAR(e) = \mathbf{R} \\ \beta_t &= \boldsymbol{\mu} + \mathbf{F}\beta_{t-1} + v_t, VAR(v) = \mathbf{Q} \end{aligned}$$

- Assuming normal error terms $\rightarrow$

$$y_t \backslash x_t \sim N\left(\mathbf{x}\beta_t, xP_{t\backslash t-1}x' + R\right)$$

- The likelihood function for each observation is

$$(2\pi)^{-n/2}\left|xP_{t\backslash t-1}x' + R\right|^{-0.5}$$

$$\exp\left(-0.5\left(y_t - x\beta_t\right)'\left(xP_{t\backslash t-1}x' + R\right)^{-1}\left(y_t - x\beta_t\right)\right)$$

$$or\ \ (2\pi)^{-n/2}\left|f_{t\backslash t-1}\right|^{-0.5}\exp\left(-0.5\eta'_{t\backslash t-1}\left(f_{t\backslash t-1}\right)^{-1}\eta_{t\backslash t-1}\right)$$

Maximum Likelihood estimation

- R, Q, μ and F are generally not known but need to be estimated
- The Kalman filter provides us with an estimate of the likelihood function of the model
- Recall

$$\begin{aligned} y_t &= \mathbf{x}\beta_t + e_t, VAR(e) = \mathbf{R} \\ \beta_t &= \boldsymbol{\mu} + \mathbf{F}\beta_{t-1} + v_t, VAR(v) = \mathbf{Q} \end{aligned}$$

- Assuming normal error terms $\rightarrow$

$$y_t \backslash x_t \sim N\left(\mathbf{x}\beta_t, xP_{t\backslash t-1}x' + R\right)$$

- The likelihood function for each observation is

$$(2\pi)^{-n/2}\left|xP_{t\backslash t-1}x' + R\right|^{-0.5} \exp\left(-0.5\left(y_t - x\beta_t\right)'\left(xP_{t\backslash t-1}x' + R\right)^{-1}\left(y_t - x\beta_t\right)\right)$$

$$or \;\; (2\pi)^{-n/2}\left|f_{t\backslash t-1}\right|^{-0.5} \exp\left(-0.5\eta'_{t\backslash t-1}\left(f_{t\backslash t-1}\right)^{-1}\eta_{t\backslash t-1}\right)$$

Maximum Likelihood estimation

Starting Values (time 0)

$$\beta_{0\backslash 0}, P_{0\backslash 0}$$

↓

Predict States (time 1...onwards)

$$\beta_{t\backslash t-1} = \mu + F\beta_{t-1\backslash t-1}$$
$$P_{t\backslash t-1} = FP_{t-1\backslash t-1}F' + Q$$

↓

Calculate Prediction Error

$$\eta_{t\backslash t-1} = y_t - x_t\beta_{t\backslash t-1}$$
$$f_{t\backslash t-1} = x_t P_{t\backslash t-1} x_t' + R$$

↓

Update States

$$K = P_{t\backslash t-1} x_t' f_{t\backslash t-1}^{-1}$$
$$\beta_{t\backslash t} = \beta_{t\backslash t-1} + K\eta_{t\backslash t-1}$$
$$P_{t\backslash t} = P_{t\backslash t-1} - Kx_t P_{t\backslash t-1}$$

↓

Calculate Likelihood

$$L_t = (2\pi)^{-n/2} |f_{t\backslash t-1}|^{-0.5} \exp\left(-0.5\eta_{t\backslash t-1}' (f_{t\backslash t-1})^{-1} \eta_{t\backslash t-1}\right)$$

- maximise $\sum_{t=1}^{T} L_t$ wrt R, Q, μ, F, A

Further Reading

- Hamilton, J.D. (1994). Time series analysis. Princeton: Princeton University Press. [Chapter 13]
- Kim, C.-J. and C. R. Nelson (1999). State-space models with regime switching. Cambridge, Massachusetts: MIT Press. [Chapter 3]

Further Reading

- Hamilton, J.D. (1994). Time series analysis. Princeton: Princeton University Press. [Chapter 13]
- Kim, C.-J. and C. R. Nelson (1999). State-space models with regime switching. Cambridge, Massachusetts: MIT Press. [Chapter 3]