State space models and Kalman filter

Business Cycles Theory

Jaromír Baxa
IES FSV UK

Introduction

- Many dynamic models of time series can be generalized into the state space models; state space models allow to incorporate information about a structure of an underlying dynamics.
- General idea: dynamics of observed (multiple) time series $y_1, \ldots, y_T$ depends on unobserved state ξ_t which is driven by stochastic process. Examples: potential GDP, common factor, volatility patterns...
- Observation (measurement) equation:

$$y_t = \boldsymbol{H}_t \xi_t + \mu_t$$

- Transition equation

$$\xi_t = \boldsymbol{F}_{t-1} \xi_{t-1} + v_t$$

- Observation and transition equations form the *state space model*.
- Matrices $\boldsymbol{H}_t$ and $\boldsymbol{F}_t$ are supposed to be known at time *t,* and can be both constant or time-varying. Both equations can be extended for exogenous variables as well.
- **Kalman filter**: Method how to recover the unobserved state when the observations are known.
- In economics, we usually do not know the matrices of coefficients and those have to be estimated as well. This is done via maximum likelihood or Bayesian techniques.

Example: Local level model

- The original time series can be decomposed into its (unobserved) components: trend defined as a random walk + noise with temporary effects.

MEASUREMENT EQUATION:

$$y_t = \beta_t + e_t \qquad e_t \sim N(0, R)$$

TRANSITION (STATE) EQUATION:

$$\beta_t = \beta_{t-1} + v_t \qquad v_t \sim N(0, Q)$$

Example: Local level model

- The original time series can be decomposed into trend and cycle assuming no noise:

MEASUREMENT EQUATION:

$$y_t = \omega_t + \tau_t$$

TRANSITION (STATE) EQUATIONS:

$$\begin{aligned} \omega_t = c + \rho_1 \omega_{t-1} + \rho_2 \omega_{t-2} + v_{1,t} \qquad & v_{1,t} \sim N(0, Q_1) \\ \tau_t = \tau_{t-1} + v_{2,t} \qquad & v_{2,t} \sim N(0, Q_2) \end{aligned}$$

Example: Local trend model

Linear growth model with random walk around linear growth and with varying drift.

MEASUREMENT EQUATION:

$$y_t = \beta_t + e_t \qquad e_t \sim N(0, R)$$

TRANSITION (STATE) EQUATIONS:

$$\begin{aligned} \beta_t &= \beta_{t-1} + \mu_{t-1} + v_{1,t} & v_{1,t} &\sim N(0, Q_1) \\ \mu_t &= \mu_{t-1} + v_{2,t} & v_{2,t} &\sim N(0, Q_2) \end{aligned}$$

State space models and Kalman filter

- Suppose a state space model with known parameters:

  $y_t = \boldsymbol{H}_t \xi_t + \mu_t$

  $\xi_t = \boldsymbol{F}_{t\text{-}1} \xi_{t\text{-}1} + v_t$

- To obtain values of the unobserved state, we use the Kalman filter: Iterative algorithm based on matrix operations that sequentially recovers the trajectory of the unobservable variable.

  1. Initialization (initial conditions for the state) $\rightarrow$

  2. Prediction (prediction of state and observation at time *t* based on information available at *t*-1)

  $$\xi_{t|t-1} = E\left(x_t | \Omega_{t-1}\right) \qquad P_{t|t-1} = E\left(\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)'\right) \qquad \rightarrow$$

  3. Updating prediction after the information about observed variable at time *t* arrives.

  (Steps 2 and 3 are repeated for all observations up to *T*.) $\rightarrow$

  4. Smoothing: utilizing the full sample prediction of the state to smooth the state.

- If parameters unknown, they are estimated using ML or Bayesian methods.

Kalman filter algorithms

$$y_t = H\xi_t + Ax_t + \mu_t$$
$$\xi_t = F\xi_{t-1} + Bx_{t-1} + \nu_t$$
$$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t \nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$$

Main steps: Initialization, Prediction, Updating, Smoothing.

0. Initialization $\qquad \xi_{t0} = \xi_0, P_0 = \Sigma_0$

Kalman filter algorithms

$$y_t = H\xi_t + Ax_t + \mu_t$$
$$\xi_t = F\xi_{t-1} + Bx_{t-1} + \nu_t$$
$$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t \nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$$

Main steps: Initialization, Prediction, Updating, Smoothing.

0. Initialization $\xi_{t0} = \xi_0, P_0 = \Sigma_0$

1. Predicting

Based on information available at $t-1$ the state $\xi_{t|t-1}$, the observed variable $y_{t|t-1}$ and covariance matrix of state variable $P_{t|t-1}$ are derived:

Prediction of state: $\xi_{t|t-1} = F\xi_{t-1|t-1} + Bx_{t-1}$

Prediction of variance of state $P_{t|t-1} = FP_{t-1|t-1}F' + Q$

Prediction of observations $y_{t|t-1} = H\xi_{t|t-1} + Ax_{t|t-1}$

Predictions of state and of variance directly from the state equation.
Prediction error P depends on previous prediction error, scaled by matrix F, and Q.
Prediction of observation from the measurement equation.

Kalman filter algorithms

$$y_t = H\,\xi_t + A\,x_t + \mu_t$$
$$\xi_t = F\,\xi_{t-1} + B\,x_{t-1} + \nu_t$$
$$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t\,\nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$$

Main steps: Initialization, Prediction, Updating, Smoothing.

0. Initialization $\quad \xi_{t0} = \xi_0,\, P_0 = \Sigma_0$

1. Predicting

Based on information available at $t-1$ the state $\xi_{t|t-1}$, the observed variable $y_{t|t-1}$ and covariance matrix of state variable $P_{t|t-1}$ are derived:

Prediction of state: $\quad \xi_{t|t-1} = F\,\xi_{t-1|t-1} + B\,x_{t-1}$

Prediction of variance of state $\quad P_{t|t-1} = F\,P_{t-1|t-1}\,F' + Q$

Prediction of observations $\quad y_{t|t-1} = H\,\xi_{t|t-1} + A\,x_{t|t-1}$

2. Updating

Inference about ξ_t is updated with a help of a new observation of y_t:

Prediction error of observations $\quad \varepsilon_t = y_t - y_{t|t-1}$

Prediction of variance of observations $\quad \psi_t = H\,P_{t|t-1}\,H' + R$

Update of state $\quad \xi_{t|t} = \xi_{t|t-1} + K_t\,\varepsilon_t$

where K_t is the Kalman gain: $\quad K_t = P_{t|t-1}\,H'(\psi)^{-1}$

Finally, update of variance of state $\quad P_{t|t} = (I - K_t H)\,P_{t|t-1}$

Updates the state for the observation at time *t*, i.e. for (part) of the error between predicted and observed realization of the observable variable y_t. The part, for which is corrected, is the Kalman gain.

Kalman filter algorithms

$y_t = H\,\xi_t + A\,x_t + \mu_t$
$\xi_t = F\,\xi_{t-1} + B\,x_{t-1} + \nu_t$
$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t\, \nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$

Main steps: Initialization, Prediction, Updating, Smoothing.

0. Initialization $\xi_{t0} = \xi_0,\, P_0 = \Sigma_0$

1. Predicting

Based on information available at $t-1$ the state $\xi_{t|t-1}$, the observed variable $y_{t|t-1}$ and covariance matrix of state variable $P_{t|t-1}$ are derived:

Prediction of state: $\xi_{t|t-1} = F\,\xi_{t-1|t-1} + B\,x_{t-1}$

Prediction of variance of state $P_{t|t-1} = F\,P_{t-1|t-1}\,F' + Q$

Prediction of observations $y_{t|t-1} = H\,\xi_{t|t-1} + A\,x_{t|t-1}$

Updates the state for the observation at time *t*, i.e. for (part) of the error between predicted and observed realization of the observable variable y_t. The part, for which is corrected, is the Kalman gain.

2. Updating

Inference about ξ_t is updated with a help of a new observation of y_t:

Prediction error of observations $\varepsilon_t = y_t - y_{t|t-1}$

Prediction of variance of observations $\psi_t = H\,P_{t|t-1}\,H' + R$

Update of state $\xi_{t|t} = \xi_{t|t-1} + K_t\,\varepsilon_t$

where K_t is the Kalman gain: $K_t = P_{t|t-1}\,H'(\psi)^{-1}$

Finally, update of variance of state $P_{t|t} = (I - K_t H)\,P_{t|t-1}$

Kalman gain depends on variance of state P (+) and on variance of observations Ψ(inversely).

If the variance of observations Ψ is small (small noise), then the"gain" shall be high and vice versa.

Kalman filter algorithms

Main steps: Initialization, predicting, updating and smoothing.

$$y_t = H\,\xi_t + A\,x_t + \mu_t$$
$$\xi_t = F\,\xi_{t-1} + B\,x_t + \nu_t$$
$$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t \nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$$

3. Smoothing

At this stage, the estimated states are revised according to the information available from the whole sample. Hence, states are updated backwards.

$$\xi_{t,T} = \xi_{t|t} + V_t\left(\xi_{t+1|T} - \xi_{t+1|t}\right)$$
$$P_{t|T} = P_{t|t} + V_t\left(P_{t+1|T} - P_{t+1|t}\right)V'_t$$

where $V_t = P_{t|t}\,F'\,P_{t+1|t}^{-1}$ is the Kalman smoothing matrix.

Kalman filter algorithms

Main steps: Initialization, predicting, updating and smoothing.

$$y_t = H\xi_t + Ax_t + \mu_t$$
$$\xi_t = F\xi_{t-1} + Bx_t + \nu_t$$
$$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t \nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$$

3. Smoothing

At this stage, the estimated states are revised according to the information available from the whole sample. Hence, states are updated backwards.

$$\xi_{t,T} = \xi_{t|t} + V_t(\xi_{t+1|T} - \xi_{t+1|t})$$
$$P_{t|T} = P_{t|t} + V_t(P_{t+1|T} - P_{t+1|t})V'_t$$

where $V_t = P_{t|t} F' P_{t+1|t}^{-1}$ is the Kalman smoothing matrix.

The smoother augments the state by unexpected difference between expected and final estimate of the state.

Kalman filter algorithms

Main steps: Initialization, predicting, updating and smoothing.

$$y_t = H\xi_t + Ax_t + \mu_t$$
$$\xi_t = F\xi_{t-1} + Bx_t + \nu_t$$
$$Cov(\mu_t \mu_t) = R, \quad Cov(\nu_t \nu_t) = Q, \quad Cov(\mu_t, \nu_t) = 0$$

3. Smoothing

At this stage, the estimated states are revised according to the information available from the whole sample. Hence, states are updated backwards.

$$\xi_{t,T} = \xi_{t|t} + V_t(\xi_{t+1|T} - \xi_{t+1|t})$$
$$P_{t|T} = P_{t|t} + V_t(P_{t+1|T} - P_{t+1|t})V'_t$$

where $V_t = P_{t|t} F' P_{t+1|t}^{-1}$ is the Kalman smoothing matrix.

The smoother augments the state by unexpected difference between expected and final estimate of the state.

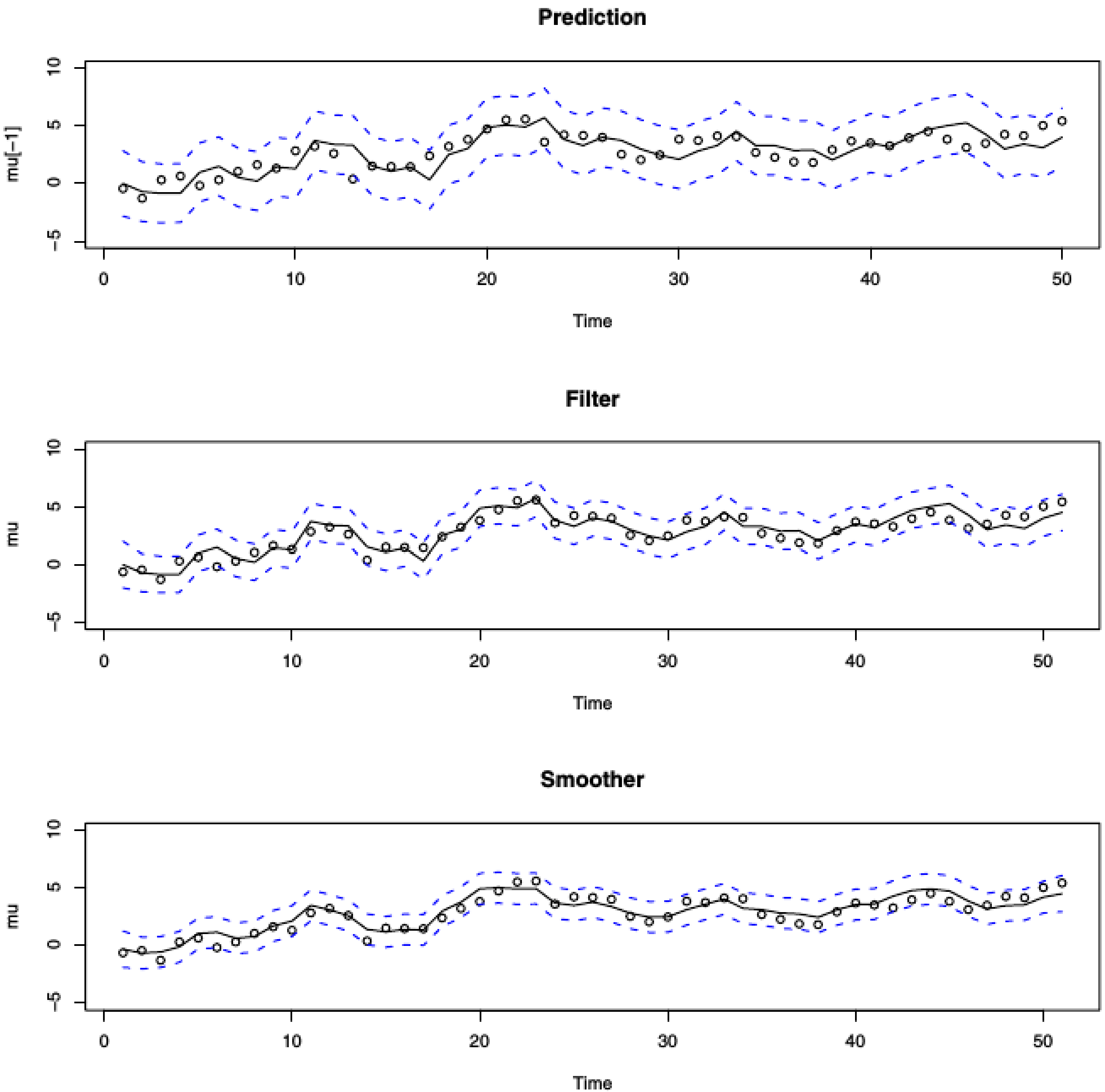
Note: The updating iterations lead to estimate of the state $\xi_{T|T}$.

Then, we use the information about the size of the last updating of the state ξ between T-1 and T to re-calculate ξ_{T-1}:

$$\xi_{T-1|T} = \xi_{T-1|T-1} + V_{T-1}(\xi_{T|T} - \xi_{T|T-1})$$

where $(\xi_{T|T} - \xi_{T|T-1})$ is the size of the last updating step.

Kalman filter: Illustration



Estimation of State space models

- Maximum likelihood.
- The conditional density function (related to information *Ω* available at time *t-1* is:

$$f\left(y_t|\Omega_{t-1}\right)=\left(2\pi\right)^{-n/2}|\psi_t|^{-1/2}\exp\left(-\frac{\varepsilon_t^{'}\psi_t\varepsilon_t}{2}\right)$$

- The log-likehood is maximized:

$$max\sum_{t=1}^{T}\log f\left(y_t|\Omega_{t-1}\right)=-\frac{nT}{2}\log\left(2\pi\right)-\frac{1}{2}\sum_{t=1}^{T}\log|\psi_t|-\frac{1}{2}\sum_{t=1}^{T}\left(\varepsilon_t^{'}\psi_t\varepsilon_t\right)$$

- Implementation: Parameters => Kalman filter => Likelihood
  - Kalman recursions for given set of parameters returns likelihood. Then, using numerical method, the neighbourhood of this set is explored to find combination with higher likelihood, and so on, until the numerical optimization converged.
- Caveat: State space models usually not identified and the assumption of monotonicity of the likelihood function does not hold quite often.
- Initialization of state space models seems to matter especially with non-stationary variables. Possible initializations: mean and variance from OLS, approximation of an unobservable variable by an observable one; or to resort to Bayesian estimation and use priors.

Applications: AR(2) model

AR2 model.

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

Observation eq.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}$$

State equation

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \underline{\quad} & \underline{\quad} \\ \underline{\quad} & \underline{\quad} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix}$$

Applications: AR(2) model

AR2 model.

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

Observation eq.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}$$

State equation

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \\ \end{pmatrix}$$

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ & \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \alpha \\ \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \end{pmatrix}$$

Applications: AR(2) model

AR2 model.

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

Observation eq.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}$$

State equation

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \_ & \_ \\ \_ & \_ \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \\ \end{pmatrix}$$

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}$$

Applications: AR(p) model

AR(p) model. $\quad y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t$

Observation eq. $\quad y_t = H X_t \quad$ with $\quad X_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}, \quad H = (1\ 0 \cdots 0)$

$\Longrightarrow y_t = y_t$

State equation

Applications: AR(p) model

ARp model. $$y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

Observation eq. $$y_t = H X_t \quad \text{with} \quad X_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}, \quad H = (1\ 0 \cdots 0)$$

$$\Longrightarrow y_t = y_t$$

Companion matrix F

$$F = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots\ 1 & 0 \end{pmatrix}$$

Applications: AR(p) model

AR~~p~~ model. $y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t$

Observation eq. $y_t = H X_t$ with $X_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}$, $H = (1\ 0 \dots 0)$

$\Longrightarrow y_t = y_t$

Companion matrix F

$$F = \begin{pmatrix} \phi_1 & \phi_2 & \dots & \phi_p \\ 1 & 0 & & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots\ 1 & 0 \end{pmatrix}$$

Transition equation $X_t = F X_{t-1} + c + \eta$ with $c = \begin{pmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\eta = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

$\Longrightarrow$ Any AR(p) $\longrightarrow$ AR(1)

Applications: ARMA(1,1) model

ARMA model.

$$y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1} \quad \ldots \quad \text{ARMA}(1,1)$$

Applications: ARMA(1,1) model

ARMA model.

$$y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1} \quad \dots \quad \text{ARMA}(1,1)$$

$$\rightarrow x_t = \begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix}$$

Observation eq.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix} + \mu$$

Applications: ARMA(1,1) model

ARMA model.

$$y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1} \quad \dots \quad \text{ARMA}(1,1)$$

$$\rightarrow x_t = \begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix}$$

Observation eq.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix} + \mu$$

Transition equation

$$\begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix} = \begin{pmatrix} \quad \end{pmatrix} \begin{pmatrix} y_{t-1} - \mu \\ \theta\varepsilon_{t-1} \end{pmatrix} + \begin{pmatrix} \quad \end{pmatrix} \quad$$

Applications: ARMA(1,1) model

ARMA model.

$$y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1} \quad \dots \quad \text{ARMA}(1,1)$$

$$\rightarrow x_t = \begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix}$$

Observation eq.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix} + \mu$$

Transition equation

$$\begin{pmatrix} y_t - \mu \\ \theta\varepsilon_t \end{pmatrix} = \begin{pmatrix} \phi & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} - \mu \\ \theta\varepsilon_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ \theta \end{pmatrix} \varepsilon_t$$

$$\rightarrow x_t = \Phi x_{t-1} + \varepsilon_t \quad \Rightarrow \text{AR}(1) \text{ again.}$$

Time-varying parameter models

- Regression models where parameters are allowed to evolve over time to track potential structural changes or policy shifts.
- Measurement equation:

$$y_t = x_t\beta_t + e_t \quad e_t \sim N(0,R)$$

- Transition equation:

$$\beta_t = \beta_{t\text{-}1} + v_t \quad v_t \sim N(0,Q)$$

- Baseline specification: coefficients follow a random walk process.
- Identification problem: portion of variance of residuals from the original regression that would be attributed to residuals of measurement equation and to variation in parameters.
- Initial values: Either OLS estimates, or training sample (several initial observations).
- Wide range of applications: time variance in monetary policy rule, coefficients of the NKPC, financial variables (time-varying CAPM betas)...

Stochastic volatility

Stochastic volatility

$y_t = \omega_t \cdots$ time varying variance

$\omega_t = \sigma_t \cdot \varepsilon_t \cdots \varepsilon_t = N(0, \sigma_\varepsilon^2)$

Stochastic volatility

Stochastic volatility

$y_t = \omega_t$ … time varying variance

$\omega_t = \sigma_t \cdot \varepsilon_t$ … $\varepsilon_t = N(0, \sigma_\varepsilon^2)$

$h_t = \log(\sigma_t^2)$ … $\boxed{h_t = c + \lambda h_{t-1} + \xi_t}$ … AR1 of $\log(\sigma_t^2)$

Observation eq.

$g = \log y^2 \to g = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} h_t \\ h_{t-1} \end{pmatrix} + k_t$

$k_t = \log(\varepsilon_t^2)$

Transition equation

$$\begin{pmatrix} h_{t+1} \\ h_t \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} h_t \\ h_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} c + \begin{pmatrix} \xi_{t+1} \\ 0 \end{pmatrix}$$

Stochastic volatility

$$y_t = \omega_t$$

In a stochastic volatility model, $h_t = \log \sigma_t^2$, the logarithm of the variance, behaves exactly as a stochastic process in the mean, such as random walks or autoregression.

$$\omega_t = \sigma_t \varepsilon_t$$

$$h_t = c + \lambda h_{t-1} + \zeta_t$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\zeta_t \sim N(0, \sigma_\zeta^2)$$

$$g_t = h_t + \kappa_t$$

where $g_t = \ln(y_t^2)$, and $\kappa_t = \ln(\varepsilon_t^2)$.

The observation equation is:

$$g_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} h_t \\ h_{t-1} \end{bmatrix} + \kappa_t$$

The state equation is:

$$\begin{bmatrix} h_{t+1} \\ h_t \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} h_t \\ h_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c + \begin{bmatrix} \zeta_{t+1} \\ 0 \end{bmatrix}$$

Dynamic factor models

- Suppose a set of k observed variables x_t that depend linearly on N unobserved common factors f_t (thus $N << k$) with idiosyncratic components u_t.

$$x_t = L f_t + u_t$$

$$f_t = A f_{t-1} + v_t$$

- Residuals are supposed to be uncorrelated.
- L contain "factor loadings".
- Factor models frequently used for coincident indices of economic activity (early application includes Stock-Watson, 1989, NBER Macroeconomic manual) or newer applications with mixed frequency data that combine quarterly data with monthly data to facilitate nowcasting.

Semi-structural models

- One of the advantages of state-space models is that they allow building and estimating models which contain explicit information about a structure of the economy.
- Such models can be a viable alternative towards the DSGE models, despite often having somewhat ad-hoc strucuture.
- Two applications:
  - Laubach-Williams: *Measuring the Natural Rate of Interest* (RES, 2003)
  - Jarocinski-Lenza: *An Inflation-Predicting Measure of the Output Gap in the Euro Area* (JMCB, 2018).

Laubach – Williams (2003)

Laubach-Williams: Measuring the Natural Rate of Interest (RES, 2003).

- Natural rate of interest = real rate of interest consistent with output at its potential and stable inflation. Central role in macro theory and monetary policy – however unobserved.
- Aim of the paper: estimate the natural rate of interest and output trend growth, as both are closely related.
- Observation equations: natural rate r^* linked to trend growth rate g and other variables z. IS curve, Phillips curve, equation linking hours h and output gap.
- Implication: Natural rate of interest gradually decreased over time, but its estimates very imprecise (s.d. around +/- 2 p.p.!). Decrease of natural rate accelerated after 2008 crisis and this movement robust across countries (Holston, Laubach and Williams, 2017)

Observation equations

1a) $r_t^* = c\, g_t + z_t$

1b) $y_t^{\text{gap}} = a_{y,1} y_{t-1}^{\text{gap}} + a_{y,2} y_{t-2}^{\text{gap}} + \frac{a_r}{2} \sum_{j=1}^{2} (r_{t-j} - r_{t-j}^*) + \varepsilon_{1,t}$

1c) $\pi_t = B_\pi(L) \pi_{t-1} + b_y y_{t-1}^{\text{gap}} + b_x x_t + \varepsilon_{2t}$

1d) $h_t = f_1 y_{t-1}^{\text{gap}} + f_2 y_{t-2}^{\text{gap}} + f_3 h_{t-1} + \varepsilon_{3t}$

Transition equations

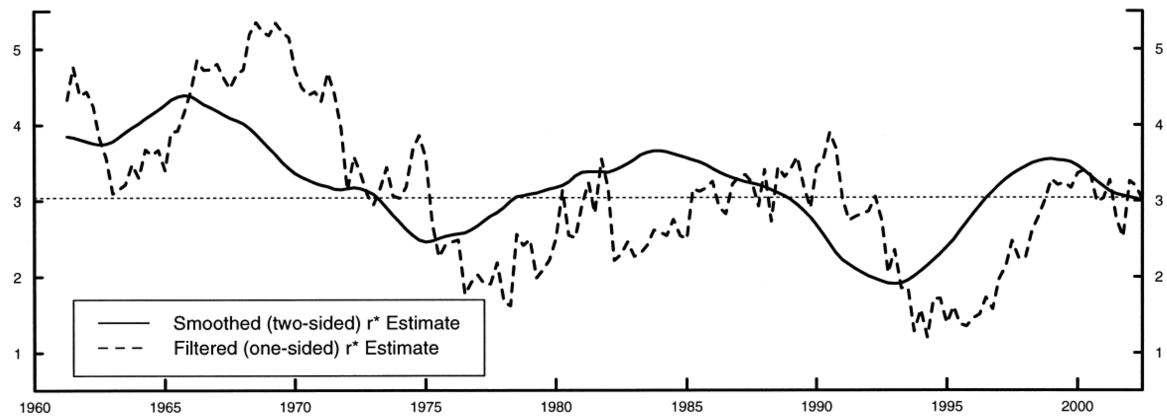
2a) $z_t = D(L) z_{t-1} + \varepsilon_{4t}$

2b) $y_t^* = y_{t-1}^* + g_{t-1} + \varepsilon_{5t}$

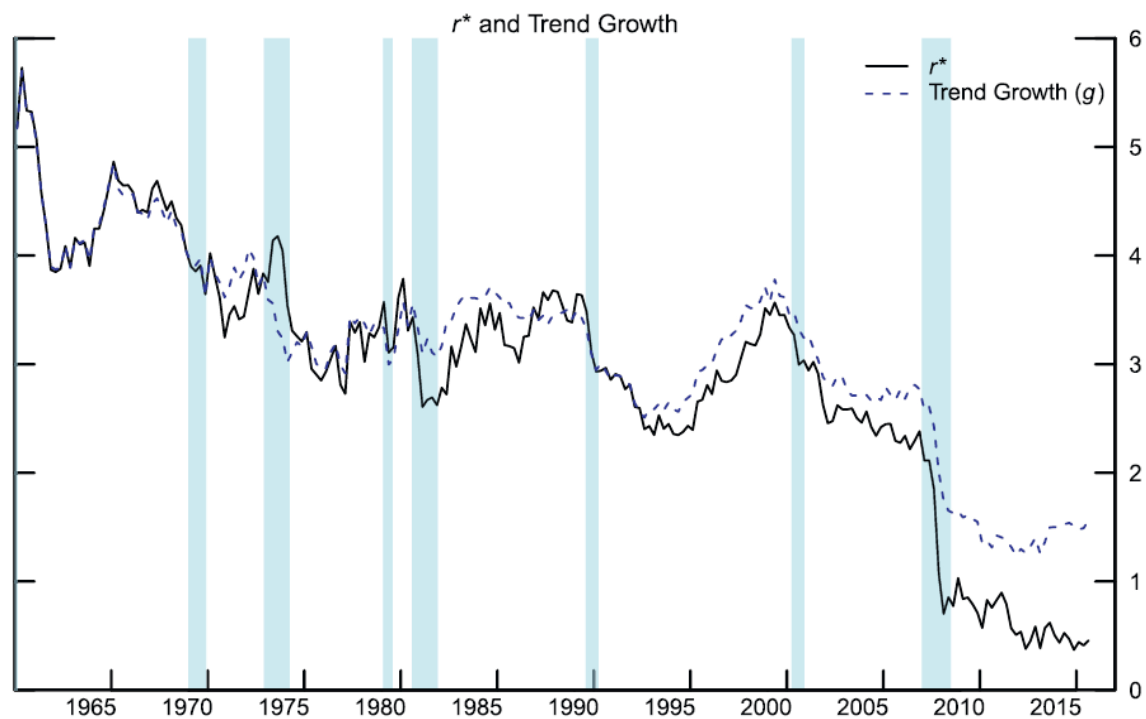
2c) $g_t = g_{t-1} + \varepsilon_{6t}$

Laubach – Williams (2003)
Holston, Laubach and Williams (2017)

FIGURE 1.—ONE- VS. TWO-SIDED ESTIMATES OF THE NATURAL RATE OF INTEREST (BASELINE MODEL)



The solid line shows the smoothed (two-sided) estimates of the natural rate of interest for the baseline specification in which z is assumed to follow a random walk. The dashed line shows the corresponding filter (one-sided) estimates.



Jarocinski - Lenza (2018)

Jarocinski-Lenza: *An Inflation-Predicting Measure of the Output Gap in the Euro Area* (JMCB, 2018).

- Aim of the paper: to provide an estimate of output gap that would be consistent with a dynamics of inflation, since the conventional estimates of output gap had a very limited power to predict inflation. Therefore, the inflation was taken as given, and the output gap derived from Phillips curve was calculated.
- Structure of the model: several real activity indicators y^N, inflation, inflation expectations => mapped to unobserved states: common stationary component g – a new indicator of output gap, trends of real activity indicators w_t and trend inflation z_t.

Observation equations

1a) $y_t^n = b^n(L) g_t + w_t^n + \varepsilon_t^n$

1b) $\pi_t - z_t = a_g(L) g_t + a_p(L)(\pi_t - 1 - z_t - 1) + \varepsilon_t^\pi$

1c) $\pi_t^e = c_0 + c_1 z_t + \varepsilon_t^e$

Transition equations

2a) $g_t = \phi_1 g_{t-1} + \phi_2 g_{t-2} + \nu_t^g$

2b) $w_t^n = d^n + w_{t-1}^n + \nu_t^n$

2c) $z_t = z_{t-1} + \nu_t^z$

- The paper contains information about priors.
- Implications: The output gap estimation in the Great Recession could have been much larger than assessed from the alternative models, therefore, the monetary policy could have been much more aggressive, too.

Jarocinski - Lenza (2018)

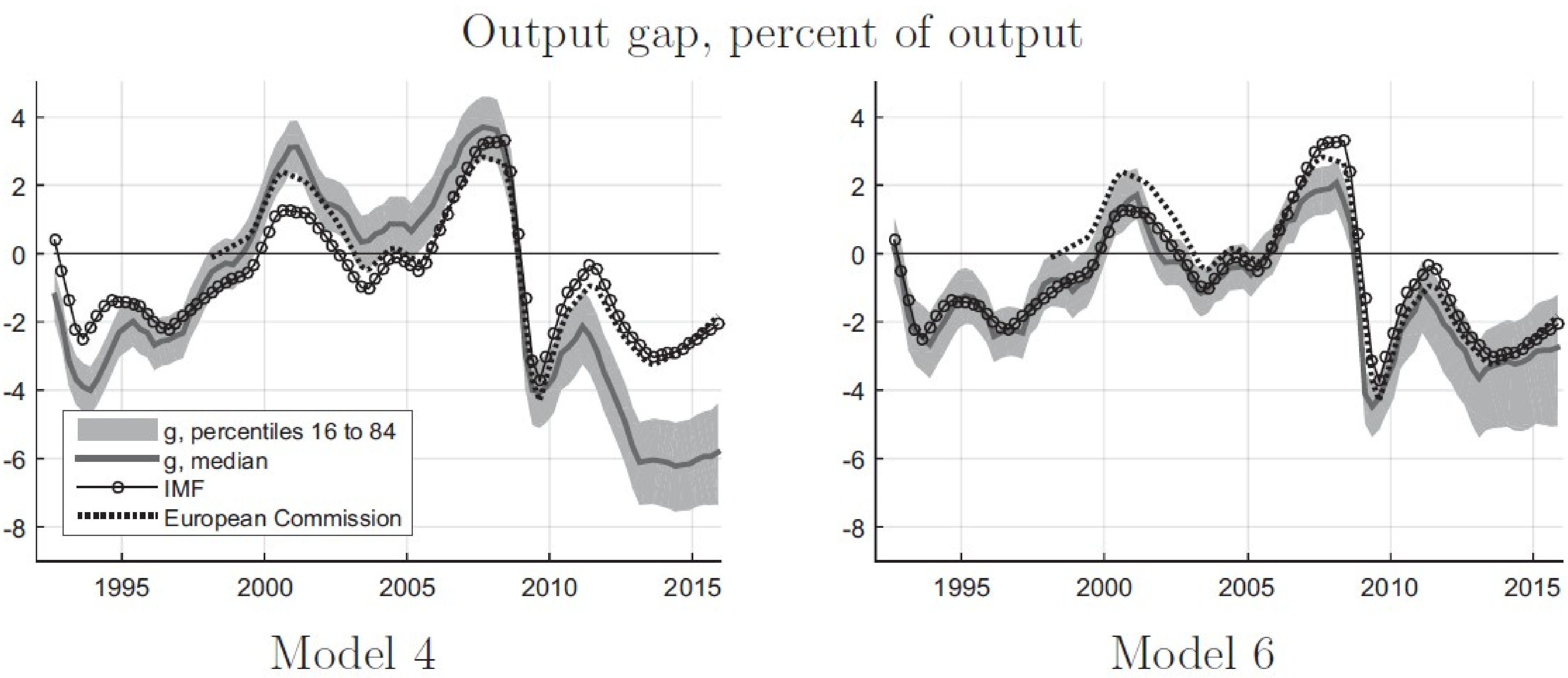


FIG. 10. Trend Output and Output Gap According to Models 4, 6, IMF and European Commission.

Summary

- The State space model: dealing with unobservables or time-varying coefficients
- Solved using Kalman filter
- Estimated using ML/Bayesian techniques
- Wide range of applications

Literature

- The principal references are:

Kalman, R.E. (1960). "A new approach to linear filtering and prediction problems". Journal of Basic Engineering. 82 (1).

Kalman, R.E.; Bucy, R.S. (1961). "New Results in Linear Filtering and Prediction Theory".

Harvey, A.C. (1990). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.

- This lecture was based on two textbooks:

Lütkepohl, H. (2005): New Introduction to Multiple Time Series Analysis. Springer (Chapter 18). Contains line-by-line derivations and proof of Kalman recursions (18.3)

Wang, P. (2009): Financial Econometrics, 2nd ed. Routledge. (Chapter 9). Accessible and straightforward introduction. Minor typo in eq. 4 of the text – exogenous variable shall be included as well.

+ lecture notes by Kevin Kotze:

https://kevinkotze.github.io/ts-4-state-space/ and https://kevinkotze.github.io/ts-4-tut/

R packages

- Interesting materials related to estimation of state space models using R in Journal of Statistical Software (code and data included)

  https://www.jstatsoft.org/article/view/v041i04

  https://www.jstatsoft.org/article/view/v036i12

- Recipes for State Space Models in R Paul Teetor: https://quantdevel.com/StateSpaceModels/index.pdf
- Structural time series models (local level...): StructTS function in base package. dlm package for more options and general state space models.
- Bayesian estimation of structural time series: bsts
- https://rstudio-pubs-static.s3.amazonaws.com/257314_131e2c97e7e249448ca32e555c9247c6.html
- http://oliviayu.github.io/post/2019-03-21-bsts/
- Multivariate structural time series: mbts
- https://arxiv.org/pdf/2106.14045
- Replication of Laubach-Williams in R https://github.com/JannesRed/rStar