Mapping the Danish Business Cycle

Advanced Empirical Macroeconomic Analysis

Martin A. Kildemark*

December 2024

*University of Copenhagen

1 Model

Obervations y_t are given by:

$$y_t = u_t + \psi_{n,t} + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma_\epsilon)$$

The stochastic trend is an integrated random walk:

$$\begin{aligned} u_t &= u_{t-1} + \beta_{t-1} \\ \beta_t &= \beta_{t-1} + \xi_t, \quad \xi_t \sim NID(0, \sigma_\xi) \end{aligned}$$

The n^{th} order cycle is given by:

$$\begin{pmatrix} \psi_{n,t} \\ \psi^*_{n,t} \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \begin{pmatrix} \psi_{n,t-1} \\ \psi^*_{n,t-1} \end{pmatrix} + \begin{pmatrix} \psi_{n-1,t-1} \\ \psi^*_{n-1,t-1} \end{pmatrix}$$

$$\begin{pmatrix} \psi_{1,t} \\ \psi^*_{1,t} \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \begin{pmatrix} \psi_{1,t-1} \\ \psi^*_{1,t-1} \end{pmatrix} + \begin{pmatrix} \kappa_t \\ \kappa^*_t \end{pmatrix}$$

with $\kappa_t \sim NID(0, \sigma_\kappa)$ and $\kappa^*_t \sim NID(0, \sigma_\kappa)$

1.1 The state space form

The model can be written in the state space form with the state vector α (following Harvey 1989):

$$\begin{aligned} y_t = Z\alpha_t + \epsilon_t, &\quad Var(\epsilon_t) = H \\ \alpha_t = T\alpha_{t-1} + R\nu_t, &\quad Var(\nu_t) = Q \end{aligned}$$

Matrices/vectors/numbers Z, H, T, R and Q are referred to as system matrices. They are assumed non-stochastic. Putting the model in state space form with a second order cycle yields the following system matrices:

$$Z = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

The transition matrix T is:

$$T = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho\cos\lambda_c & \rho\sin\lambda_c & 0 & 0 \\ 0 & 0 & -\rho\sin\lambda_c & \rho\cos\lambda_c & 0 & 0 \\ 0 & 0 & 1 & 0 & \rho\cos\lambda_c & \rho\sin\lambda_c \\ 0 & 0 & 0 & 1 & -\rho\sin\lambda_c & \rho\cos\lambda_c \end{pmatrix}$$

The matrix R that links the disturbances to the states is:

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The state vector α_t is:

$$\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \\ \psi_{1,t} \\ \psi^*_{1,t} \\ \psi_{2,t} \\ \psi^*_{2,t} \end{pmatrix}$$

1.2 The Kalman Filter

The Kalman filter provides an optimal estimator for the state space vector at time t. The estimator of α_t is denoted as a_t. P_t is the covariance matrix of the estimation error (6*6 matrix in this case):

$$P_{t-1} = E[(\alpha_{t-1} - a_{t-1})(\alpha_{t-1} - a_{t-1})']$$

Given a_{t-1} and P_{t-1}, the optimal estimate of α_t is given by the following two equations called the *prediction equations*:

$$a_{t|t-1} = Ta_{t-1}$$
$$P_{t|t-1} = TP_{t-1}T' + RQR'$$

When the observation of y_t becomes available the estimator is updated through the *updating equations*:

$$\begin{aligned} a_t &= a_{t|t-1} + P_{t|t-1}Z'F_t^{-1}(y_t - Za_{t|t-1}) \\ P_t &= P_{t|t-1} - P_{t|t-1}Z'F_t^{-1}ZP_{t|t-1} \end{aligned}$$

where

$$F_t = ZP_{t|t-1}Z' + H$$

1.3 The Kalman Smoother

The Kalman filter provides optimal estimates of the state vector at time t, given the information available at time t. More precise estimates can be obtained by conditioning on the full dataset, so both data observed at time t, and data from all following periods. This can be done with a *Kalman smoother*. There are multiple algorithms to obtain the smoothed estimates. I apply the *fixed-interval smoothing* algorithm, which is run after the Kalman filter using stored estimates of a_t and P_t. The algorithm is a backwards recursion as follows:

$$a_{t|T} = a_t + P_t^*(a_{t+1|T} - Ta_t)$$
$$P_{t|T} = P_t + P_t^*(P_{t+1|T} - P_{t+1|t})P_t^{*\prime}$$

where

$$P_t^* = P_tT'P_{t+1|t}^{-1}$$

2 Maximum Likelihood

For a Gausian linear model the likelihood function can be written as (Harvey 1989):

$$logL = -\frac{NT}{2}log2\pi - \frac{1}{2}\sum_{t=1}^{T} log|F_t| - \frac{1}{2}\sum_{t=1}^{T} v_t' F_t^{-1} v_t$$

where v are prediction errors $v_t = y_t - \tilde{y}_{t|t-1}$ and F is a covariance matrix so $v_t \sim NID(0, F_t)$ (just a variance in the univariat case). These are obtained from the Kalman filter.

3 Bayesian Estimation and MCMC

When aplying a Bayesian approach estimates are derived from Bayes' theorem:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

where $p(y)$ is known as the *marginal likelihood* and given by $p(y) = \int_{\Theta} p(y|\theta)p(\theta)d\theta$. The marginal likelihood is not available analytically, so the posterioer distribution is unknown. TO estimate the posterior distribution *Markov Chain Monte Carlo (MCMC)* methods can be used. The basic idea behind MCMC is to draw an initial set of parameters. Then you take a step in the Markov Chain by drawing a new set of parameters based only on the current parameters. Having drawn the new parameters, you choose whether to move to the new parameters or stay put by comparing the probabilities of the parameters and using an acceptance criterion. Then given the current parameters new parameters are drawn and the process continues. After a large amount of steps the distribution of the accepted parameters will follow the true posterior distribution. To generate the draws in the chain I follow Harvey et al. (2007) and implement a *Metropolis-Hastings within Gibbs* sampling algorithm, which entails drawing some parameters from a Gibbs algorithm while other parameters are drawn from a Metropolis-Hastings algorithm.

3.1 Metropolis-Hastings Sampling

3.2 Gibbs Sampling

3.3 Implementation

3.3.1 Initialization (Metropolis-Hastigns)

For s = 1,..,10,000 draws:

Draw unbounded candidate vector Γ^* from multivariate normal distribution with mean Γ_{s-1} and variance-covariance ωI where ω is a scaling constant to get acceptance rate 25-35 pct.

New draw is accepted with probability η

$$\Gamma_s = \begin{cases} \Gamma_* & \text{with probability } \eta \\ \Gamma_{s-1} & \text{with probability } 1-\eta \end{cases}$$

where

$$\eta = min\left(1, \frac{p(y|\theta^*)p(\theta^*)J(\Gamma^*)}{p(y|\theta_{s-1})p(\theta_{s-1})J(\Gamma_{s-1})}\right)$$

Where $J(\Gamma)$ are the Jacobians of the transformed variables: $\frac{d\theta}{d\Gamma}$. In practice everything is done in logs.

First half of the initialization draws are discarded

3.3.2 Recursion

for q = 1,..., 20,000 draws: Metropolis-Hastings step to sample parameters

Set Σ to sample covariance of the chain Γ_s from the initialization step. A candidate vector for the unbounded paramters Γ^* is drawn form a multivariat distribution with mean Γ_{q-1} and variance $\omega\Sigma$ with ω being a scaling constant to get 25-35 acceptance rate. New draw is accepted as in the initialization. Discard draws from burn in period.

For $q > 10,000$: After burn-in period sample unobserved states α using Gibbs sampling as follows:

- Draw α^+ and y^+ by initializing $\alpha_1^+ \sim N(0, P_1)$ and forward recursion using the model
- Construct an artificial series $y^* = y - y^+$. Compute $\hat{\alpha}^* = E(\alpha|y^*)$ by putting y^* through the Kalman smoother
- Get $\tilde{\alpha} = \hat{\alpha}^* + \alpha^+$. Then $\tilde{\alpha}$ is a draw from $p(\alpha|\theta, y)$

3.4 Priors and Jacobians

Draws of θ are drawn from the prior distribution specified by the researcher. These priors may be bounded in their support, such as a variance, which must be positive. I follow Hazenzagel and use the following priors

Name	Support	Density	Parameter 1	Parameter 2
σ^2	$(0, \infty)$	Inverse-Gamma	3	1
ρ	$[0.001, 0.970]$	Uniform	0.001	0.970
λ	$[0.001, \pi]$	Uniform	0.001	π

Table 1: Prior distributions in Hazenzagel

To get unbounded support the following transformations are used:

$$\begin{aligned} \Gamma^{U,\beta} &= \ln\left(\frac{\theta^{U,\beta} - a}{b - \theta^{U,\beta}}\right) \\ \Gamma^{IG} &= \ln(\theta^{IG} - a) \end{aligned}$$

where a is the lower bound and b is the upper bound of their support. The inverse of the transformations are then

$$\begin{aligned} \theta^{U,\beta} &= \frac{a + bexp(\Gamma^{U,\beta})}{1 + \exp(\Gamma^{U,\beta})} \\ \theta^{IG} &= \exp(\Gamma^{IG}) + a \end{aligned}$$

and the log of the derivatives are

$$\begin{aligned} \ln\left(\frac{d\theta^{U,\beta}}{d\Gamma^{U,\beta}}\right) &= \ln(b - a) + \Gamma^{U,\beta} - 2\ln(1 + \exp(\Gamma^{U,\beta})) \\ \ln\left(\frac{d\theta^{IG}}{d\Gamma^{IG}}\right) &= \Gamma^{IG} \end{aligned}$$

4 Kim and Nelson Model

Now we consider a Markov Switching State Space Model where the economy can switch between recession, $S_t = 1$ and normal times $S_t = 0$. Obervations y_t are given by:

$$y_t = u_t + \psi_{n,t}$$

The stochastic trend is an integrated random walk:

$$\begin{aligned}
u_t &= u_{t-1} + \beta_{t-1} + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma_{\epsilon,S_t}) \\
\sigma_{\epsilon,S_t} &= \sigma_{\epsilon,0}(1 - S_t) + \sigma_{\epsilon,1} S_t \\
\beta_t &= \beta_{t-1} + \xi_t, \quad \xi_t \sim NID(0, \sigma_\xi)
\end{aligned}$$

The n^{th} order cycle is given by:

$$\begin{pmatrix} \psi_{n,t} \\ \psi^*_{n,t} \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \begin{pmatrix} \psi_{n,t-1} \\ \psi^*_{n,t-1} \end{pmatrix} + \begin{pmatrix} \psi_{n-1,t-1} \\ \psi^*_{n-1,t-1} \end{pmatrix}$$

$$\begin{pmatrix} \psi_{1,t} \\ \psi^*_{1,t} \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \begin{pmatrix} \psi_{1,t-1} \\ \psi^*_{1,t-1} \end{pmatrix} + \begin{pmatrix} \omega S_t \\ \omega S_t \end{pmatrix} + \begin{pmatrix} \kappa_t \\ \kappa^*_t \end{pmatrix}$$

with $\kappa_t \sim NID(0, \sigma_{\kappa,S_t})$ and $\kappa^*_t \sim NID(0, \sigma_{\kappa,S_t})$ where $\sigma_{\kappa,S_t} = \sigma_{\kappa,0}(1 - S_t) + \sigma_{\kappa,1} S_t$.

S_t evolves according to a first order Markov process:

$$\begin{aligned}
p(S_t = 1 | S_{t-1} = 1) = p \\
p(S_t = 0 | S_{t-1} = 0) = q
\end{aligned}$$

4.1 The Filter

Conditional on $S_{t-1} = i$ and $S_t = j$ the Kalman filter is

$$\begin{aligned}
a_{t|t-1}^{i,j} &= Ta_{t-1}^{i} + c_t^j \\
P_{t|t-1}^{i,j} &= TP_{t-1}^{i}T' + RQ^jR' \\
v_t^{i,j} &= y_t - Za_{t|t-1}^{i,j} \\
F_t^{i,j} &= ZP_{t|t-1}^{i,j}Z' + H \\
a_t^{i,j} &= a_{t|t-1}^{i} + P_{t|t-1}^{i,j}Z'(F_t^{i,j})^{-1}v_t^{i} \\
P_t^{i,j} &= P_{t|t-1}^{i,j} - P_{t|t-1}^{i,j}Z'(F_t^{i,j})^{-1}ZP_{t|t-1}^{i,j}
\end{aligned}$$

If we iterate through the filter, we have to consider the entire history of S. As S can take two values: 0 and 1, we would have to consider 2^T cases, which is not feasible. As such after every iteration we reduce the dimension of $a_t^{i,j}$ and $P_t^{i,j}$ at each iteration from 2×2 matrices to 2×1 vectors thus obtaining a_t^j and P_t^j. This is done using the following approximations:

$$a_t^j = \frac{\sum_{i=0}^{1} p(S_{t-1} = i, S_t = j|\Psi_t)a_t^{i,j}}{p(S_t = j|\Psi_t)}$$

$$P_t^j = \frac{\sum_{i=0}^{1} p(S_{t-1} = i, S_t = j|\Psi_t)\left(P_t^{i,j} + (a_t^j - a_t^{i,j})(a_t^j - a_t^{i,j})'\right)}{p(S_t = j|\Psi_t)}$$

where Ψ_t represent information available at time t. The probabilities are given by:

$$p(S_t = j|\Psi_t) = \sum_{i=0}^{1} p(S_{t-1} = i, S_t = j|\Psi_t)$$

and

$$p(S_{t-1} = i, S_t = j|\Psi_t) = \frac{p(y_t|S_{t-1} = i, S_t = j, \Psi_{t-1}) \cdot p(S_{t-1} = i, S_t = j|\Psi_{t-1})}{p(y_t|\Psi_{t-1})}$$

where

$$
\begin{aligned}
p(y_t|S_{t-1}=i, S_t=j, \Psi_{t-1}) &= \frac{1}{\sqrt{2\pi F_t^{i,j}}} exp\left(-\frac{(v_t^i)^2}{2F_t^{i,j}}\right) \\
p(y_t|\Psi_{t-1}) &= \sum_{i=0}^{1}\sum_{j=0}^{1} p(y_t, S_{t-1}=i, S_t=j|\Psi_{t-1}) \\
&= \sum_{i=0}^{1}\sum_{j=0}^{1} p(y_t|S_{t-1}=i, S_t=j, \Psi_{t-1}) \cdot p(S_{t-1}=i, S_t=j|\Psi_{t-1})
\end{aligned}
$$

and

$$p(S_{t-1}=i, S_t=j|\Psi_{t-1}) = p(S_t=j|S_{t-1}=i) \cdot p(S_{t-1}=i|\Psi_{t-1})$$

Where $p(S_{t-1}=i|\Psi_{t-1})$ is obtained in the previous iteration. In the initial iteration set $p(S_0=0|\Psi_0)=\frac{1-p}{2-p-q}$ and $p(S_0=1|\Psi_0)=\frac{1-q}{2-p-q}$. The log-likelihood is then obtained through the iterations and given by

$$logL = \sum_{t=1}^{T} log\left(p(y_t|\Psi_{t-1})\right)$$

Smoothed estimates can no longer be obtained with the Kalman smoother due to the non-linearity caused by the switching structure. As such I use the filtered state estimates and not smoothed estimates. To obtain the estimated states at time t I compute $a_t=\sum_{j=0}^{1}\sum_{i=0}^{1} p(S_{t-1}=i, S_t=j|\Psi_t)a_t^{i,j}$ when running the filter with the optimal parameters.

5 PMCMC to estimate the Kim & Nelson Model

References

Álvarez, L. J. and Gómez-Loscos, A. (2018). A menu on output gap estimation methods. *Journal of Policy Modeling*, 40(4):827–850.

Chan, J. C. (2017). Notes on bayesian macroeconometrics. *Manuscript available at http://joshuachan. org*.

Chan, J. C. and Strachan, R. W. (2023). Bayesian state space models in macroeconometrics. *Journal of Economic Surveys*, 37(1):58–75.

Harvey, A. C., Trimbur, T. M., and Van Dijk, H. K. (2007). Trends and cycles in economic time series: A bayesian approach. *Journal of Econometrics*, 140(2):618–649.