Intensive Courses in the context of the Jean Monnet Chair:

Big data in official statistics

Block 3: Structural time series models (short version)

11 DECEMBER 2019,

UNIVERSITY OF PISA

Jan van den Brakel

Statistics Netherlands and Maastricht University

Introduction

Time series models:

1. Box & Jenkens ARIMA models

2. Structural time series models

Ad. 1:

The approach followed by Box and Jenkins (1989) for modelling time series starts by making an observed series stationary. Informally spoken, this implies that the trend in an observed series is removed by taking differences between subsequent periods. Seasonal patterns are removed in a similar way by taking differences between the observations of the same quarters or months of two successive years. Once an observed series is made stationary, it is modelled with autoregressive and moving average components.

Ad. 2:

Structural time series modelling follows a more direct and intuitive approach (subjective opinion). An observed series is directly modelled without attempting to remove non-stationarity through differencing of the observed series. This is the approach followed by authors like Harvey (1989), and Durbin and Koopman (2012).

Structural Time Series Models

Observed series $y_t, \quad t = 1, \ldots, T$.

Structural Time Series (STS) models decompose an observed series in:

1. Trend (L_t)
2. Seasonal (S_t)
3. Cycles (γ_t)
4. Regression component ($\boldsymbol{\beta}_t' \mathbf{x}_t$)
5. White noise (I_t)

Additive model:

$$y_t = L_t + S_t + \gamma_t + \boldsymbol{\beta}_t' \mathbf{x}_t + I_t, \quad t = 1, \ldots, T.$$

Multiplicative model:

$$y_t = L_t \times S_t \times \gamma_t \times \boldsymbol{\beta}_t' \mathbf{x}_t \times I_t, \quad t = 1, \ldots, T.$$

(Additive again after taking logs)

Local Level Model

Very simple trend model: L_t is a random walk:

$$y_t = L_t + I_t \qquad I_t \simeq \mathcal{N}(0, \sigma_I^2)$$

$$L_t = L_{t-1} + \zeta_t \quad \zeta_t \simeq \mathcal{N}(0, \sigma_\zeta^2)$$

Note:

$$y_t = L_0 + \sum_{t=1}^{t} \zeta_t + I_t$$

- Serial correlation between observations y_t. This makes routine computations from normal regression theory inefficient
- Filtering and smoothing algorithms developed as an alternative
- Express STS model as a state space model
- Local level model is already in state space representation
- Kalman filter to obtain optimal estimates for L_t

Local Linear Trend Model

Popular trend model for economic time series:

$$y_t = L_t + I_t \qquad I_t \simeq \mathcal{N}(0, \sigma_I^2)$$

$$L_t = L_{t-1} + R_{t-1} + \zeta_t \qquad \zeta_t \simeq \mathcal{N}(0, \sigma_\zeta^2)$$

$$R_t = R_{t-1} + \tau_t \qquad \tau_t \simeq \mathcal{N}(0, \sigma_\tau^2)$$

- L_t often referred to as the level
- R_t interpreted as a slope parameter
- Trend models with random levels are often volatile

Exercise:

- What happens if for the local level model $\sigma_\zeta^2 = 0$? (Illustrate with a graph.)
- What happens if for the local linear model $\sigma_\zeta^2 = 0$ and $\sigma_\tau^2 = 0$? (Illustrate with a graph.)

Smooth Trend Model

Special case of the local linear trend model:

$$y_t = L_t + I_t \qquad I_t \simeq \mathcal{N}(0, \sigma_I^2)$$

$$L_t = L_{t-1} + R_{t-1}$$

$$R_t = R_{t-1} + \tau_t \qquad \eta_t \simeq \mathcal{N}(0, \sigma_\tau^2)$$

- Only the slope is random
- Results in more stable trend patterns

Seasonal components

- Model a cycle with a period of one year
- Models:
  - Dummy seasonal model
  - Trigonometric seasonal models

More details on modeling seasonal effects in structural time series models:

- Harvey (1989), Section 4.1: dummy seasonal and trigonometric seasonal models
- Durbin and Koopman (2012), Section 3.2: dummy seasonal and trigonometric seasonal models

State Space Representation

State space representation STS model:

1. Measurement equation: $y_t = \mathbf{Z}\boldsymbol{\alpha}_t + I_t$
   - $\boldsymbol{\alpha}_t$: vector with state variables (trend, seasonal, etc)
   - $\mathbf{Z}$: Design matrix measurement equation
   - $I_t \simeq \mathcal{N}(0, \sigma_I^2)$
2. Transition equation: $\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t$
   - $\mathbf{T}$: Design matrix transition equation
   - $\boldsymbol{\eta}_t$: vector disturbances of the state variables (trend, seasonal, etc)
   - $\boldsymbol{\eta}_t \simeq \mathcal{N}(\mathbf{0}, \mathbf{H})$

Ad. 1: The measurement equation describes how the observed series depends on unobserved state variables that describe trend, seasonal components, regression components, etc.

Ad.2: The transition equation describes how the state variables evolve over time. More precisely; how they change from one period to the next.

Exercise

Give the state space representation for the local linear trend model:

$$
\begin{aligned}
y_t &= L_t + I_t & I_t &\simeq \mathcal{N}(0, \sigma_I^2) \\
L_t &= L_{t-1} + R_{t-1} + \zeta_t & \zeta_t &\simeq \mathcal{N}(0, \sigma_\zeta^2) \\
R_t &= R_{t-1} + \tau_t & \tau_t &\simeq \mathcal{N}(0, \sigma_\tau^2)
\end{aligned}
$$

State space representation local linear trend model

- Measurement equation: $y_t = \mathbf{Z}\boldsymbol{\alpha}_t + I_t$

$$\mathbf{Z} = (1 \;\; 0)$$

$$\boldsymbol{\alpha}_t = (L_t \;\; R_t)'$$

$$\Rightarrow y_t = (1 \;\; 0) \begin{pmatrix} L_t \\ R_t \end{pmatrix} + I_t$$

$$I_t \simeq \mathcal{N}(0, \sigma_I^2)$$

- Transition equation: $\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t$

$$\mathbf{T} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\eta}_t = (\zeta_t \;\; \tau_t)'$$

$$\Rightarrow \begin{pmatrix} L_t \\ R_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} L_{t-1} \\ R_{t-1} \end{pmatrix} + \begin{pmatrix} \zeta_t \\ \tau_t \end{pmatrix}$$

$$\boldsymbol{\eta}_t \simeq \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\zeta^2 & 0 \\ 0 & \sigma_\tau^2 \end{pmatrix} \right)$$

Kalman filter

- Structural time series models in state space form
- Kalman filter to obtain optimal estimates for state variables (and signal)
- Recursive algorithm that gives optimal estimates for $\boldsymbol{\alpha}_t$ based on the information available at time t
- Kalman filter recursion runs from $t = 1, ..., T$, and gives BLUP's for $\boldsymbol{\alpha}_t$ given information obtained until period t
  - $\mathbf{a}_t$: filtered estimates for $\boldsymbol{\alpha}_t$
  - $\mathbf{P}_t$: covariance matrix of $\mathbf{a}_t$
- Kalman filter recursion:
  - Prediction equations:

$$\mathbf{a}_{t|t-1} = \mathbf{T}\mathbf{a}_{t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{T}\mathbf{P}_{t-1}\mathbf{T}' + \mathbf{H}$$

The prediction equations follow directly from the transition equation.

– From the measurement equation it follows:

$$\hat{y}_{t|t-1} = \mathbf{Z}\mathbf{a}_{t|t-1}$$

– Innovation (new information if y_t becomes available):

$$\nu_t = y_t - \hat{y}_{t|t-1} = \mathbf{Z}(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + I_t$$

– Variance innovations

$$f_t = \mathbf{Z}\mathbf{P}_{t|t-1}\mathbf{Z}' + \sigma_I^2$$

– Updating equations (BLUP for $\boldsymbol{\alpha}_t$):

$$\begin{aligned}\mathbf{a}_t &= \mathbf{a}_{t|t-1} + \frac{\nu_t}{f_t}\mathbf{P}_{t|t-1}\mathbf{Z}' \\ \mathbf{P}_t &= \mathbf{P}_{t|t-1} - \frac{1}{f_t}\mathbf{P}_{t|t-1}\mathbf{Z}'\mathbf{Z}\mathbf{P}_{t|t-1}\end{aligned}$$

The updating equations follow from the assumption that $\boldsymbol{\alpha}_0$, I_t, and $\boldsymbol{\eta}_t$ are multivariate normally distributed and subsequently the conditional distribution of $\boldsymbol{\alpha}_t$ given y_t. For a proof see Harvey (1989), Ch. 3.

- To start the filter:

  – Initial values for $\mathbf{a}_0$ and $\mathbf{P}_0$ are known

- Covariance matrices of the measurement and system equation are known, i.e. σ_I^2 and $\mathbf{H}$

Smoothing

- $\mathbf{a}_t$: filtered estimates for $\boldsymbol{\alpha}_t$ given information obtained until period t
- Smoothing improves $\mathbf{a}_t$ using information obtained after period t
- Widely applied smoothing algorithm: fixed interval smoother
- Recursive algorithm that starts with the final quantities $\mathbf{a}_T$ and $\mathbf{P}_T$ and runs back from $t = T - 1, ..., 1$
- Smoothed BLUP's of $\boldsymbol{\alpha}_t$:

$$\mathbf{a}_{t|T} = \mathbf{a}_t + \mathbf{P}_t\mathbf{T}'\mathbf{P}_{t+1|t}^{-1}(\mathbf{a}_{t+1|T} - \mathbf{T}\mathbf{a}_t)$$

- Covariance matrix of prediction errors of $\mathbf{a}_{t|T}$:

$$\mathbf{P}_{t|T} = \mathbf{P}_t + \mathbf{P}_t\mathbf{T}'\mathbf{P}_{t+1|t}^{-1}(\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t})\mathbf{P}_{t+1|t}^{-1}\mathbf{T}\mathbf{P}_t$$

Multivariate State Space Models

- Measurement equation: $\mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{i}_t$

  with $\mathbf{y}_t = (y_{1,t}, ..., y_{n,t})'$

$$\mathbf{i}_t \simeq \mathcal{N}(\mathbf{0}, \mathbf{G})$$

$$\mathbf{G} = \mathbf{Diag}(\sigma^2_{I_1}, ..., \sigma^2_{I_n})$$

- Transition equation: $\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t$

$$\boldsymbol{\eta}_t \simeq \mathcal{N}(\mathbf{0}, \mathbf{H})$$

Multivariate State Space Models

- Kalman filter recursion:
  - Prediction equations:

$$\mathbf{a}_{t|t-1} = \mathbf{T}\mathbf{a}_{t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{T}\mathbf{P}_{t-1}\mathbf{T}^{'} + \mathbf{H}$$

  - Updating equations:

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{Z}^{'}\mathbf{F}_t^{-1}(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{Z}^{'}\mathbf{F}_t^{-1}\mathbf{Z}\mathbf{P}_{t|t-1}$$

  - Covariance matrix innovations:

$$\mathbf{F}_t = \mathbf{Z}\mathbf{P}_{t|t-1}\mathbf{Z}^{'} + \mathbf{G}$$

- Required:
  - values for hyperparameters $\mathbf{G}$ and $\mathbf{H}$:

    Maximum Likelihood
  - initial values for $\mathbf{a}_0$ and $\mathbf{P}_0$:

    Diffuse or exact initialization

Starting values for the Kalman filter: $\mathbf{a}_0$ and $\mathbf{P}_0$

- Sometimes a-priori information: exact initialization
- If there is no a-priori information; distinguish between
  - Non-stationary state variables:
    - Diffuse initialization
    - $\mathbf{a}_0 = \mathbf{0}$
    - $\mathbf{P}_0 = \kappa\mathbf{I}$ with $\kappa = 10^7$
  - Stationary state variables:
    - Exact initialization
    - $\mathbf{a}_0 = \mathbf{0}$ (expected value)
    - $\mathbf{P}_0$ derived from its process
- State space model with d non stationary state variables
- First d observations are used to construct a proper distribution for the non-stationary state variables

Model evaluation

Model assumptions:

- Disturbance terms measurement and transition equations are normally and serially independent distributed
- $\Rightarrow$ Innovations or one-step forecast errors are normally and serially independent distributed

  Follows from the prediction error decomposition.
- Model diagnostics are focussed on checking the assumption that standardized innovations are standard normal distributed

Standardized innovations:

$$\tilde{\nu}_t = \frac{\nu_t}{\sqrt{f_t}}$$

with:

$$\nu_t = y_t - \hat{y}_{t|t-1}$$

$$f_t = Var(\nu_t) = \mathbf{Z}\mathbf{P}_{t|t-1}\mathbf{Z}' + \sigma_I^2$$

Recall that the first d time periods are ignored in the evaluation of the likelihood function for d diffuse state variables.

Normality:

- Bowman-Shenton test on normality
- QQ-plots
- Histogram
- Plot of $\tilde{\nu}_t$ for $t = d, \ldots, T$ with 95% confidence interval

Heteroscedasticity: F-test:

$$F = \frac{\sum_{t=d}^{h+d} \tilde{\nu}_t^2}{\sum_{t=T-h-d+1}^{T} \tilde{\nu}_t^2} \simeq F_h^h$$

F-test based on the sum over squared innovations for two exclusive subsets of the sample of equal length h.

Serial correlation:

- Autocorrelogram based on autocorrelations
- Liung Box test
- Durbin-Watson test

See Durbin and Koopman (2012) Section 2.12 and 7.5 for more details.

Model selection and comparison

Likelihood-based diagnostics:

- Akaike information criterion

$$AIC = \frac{1}{(T-d)}[-2log(L) + 2(q+p)]$$

- Bayesian information criterion

$$BIC = \frac{1}{(T-d)}[-2log(L) + log(T-d)(q+p)]$$

  q: number of hyperparameters (estimated with ML)

  p: number of state variables

  d: number of non-stationary state variables

  L: Likelihood

- Nested models: Likelihood ratio test

$$LR = 2 * [\ log(L[M_{alt}]) - log(L[M_{null}])\ ] \simeq \chi^2_r$$

  - M_{alt}: extended model under alt. hypothesis
  - M_{null}: reduced model under the null hypothesis
  - r: d.f. $\Rightarrow$ number of parameters equal to zero

- Evaluate the contribution of state variables: plots of the smoothed estimates with 95% confidence interval

See Durbin and Koopman (2012) Section 7.4 for more details.

Software

Software for STM:

- Eviews
- SAS
- R: package KFAS
- Oxmetrics:
  - STAMP
  - Ssfpack

References

Box, G. and Jenkins, G. (1989). *Time series analysis: forecasting an control*. Holden-Day, San Francisco.

Durbin, J. and Koopman, S. (2012). *Time Series Analysis by State Space Methods (second edition)*. Oxford University Press, Oxford.

Harvey, A. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge.