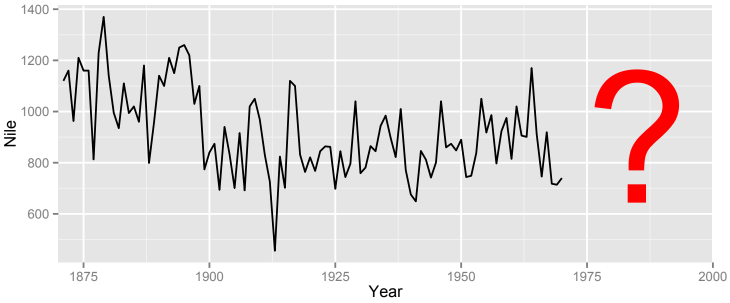
Dynamic Linear Models And Kalman Filtering

Dr. Holger Zien

12th December 2014

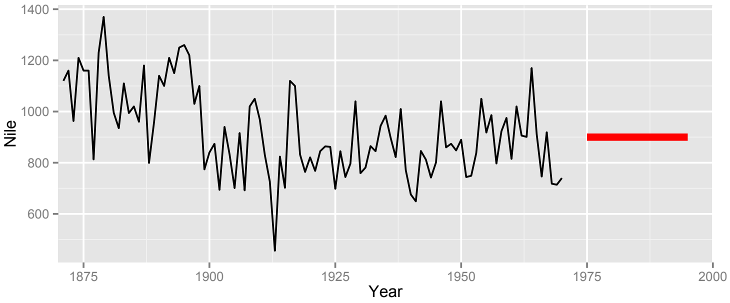
Introduction

One of The Most Common Problems Is To Forecast Time Series



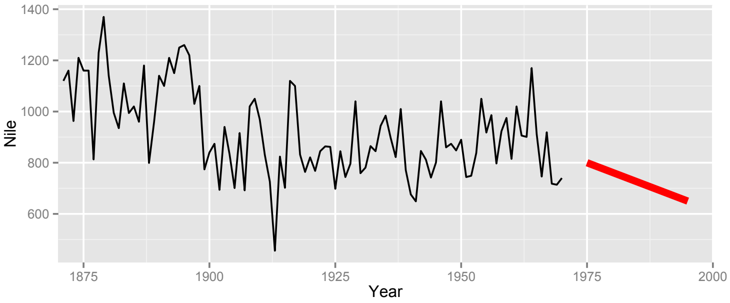
Introduction

One of The Most Common Problems Is To Forecast Time Series



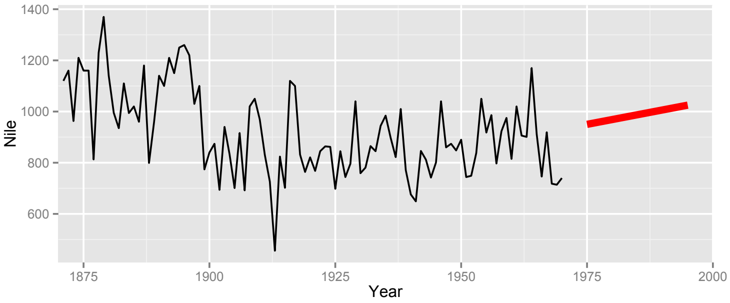
Introduction

One of The Most Common Problems Is To Forecast Time Series



Introduction

One of The Most Common Problems Is To Forecast Time Series



Introduction

Typical Time Series In Classical Textbooks

- many samples
- more or less stationary
- typical for physical measurements
- the method of choice is ARMA

Introduction

In Practice We Are Facing A Different Kind of Time Series

- ridiculous short
- often non–stationary
- ARMA models do not work
- Sometimes Bayesian modeling of Dynamic Linear Models/Kalman Filtering will help

Introduction

Autoregressive Moving Average Model (ARMA)

$$y_t = \varepsilon_t + \underbrace{\sum_{i=1}^{p} \phi_i y_{t-i}}_{\text{autoregressive part}} + \underbrace{\sum_{j=1}^{q} \psi_j \varepsilon_{t-j}}_{\text{moving average part}}$$

Dynamic Linear Model

Simplest Version

observation eq.	$y_t = F\theta_t + \nu_t$
system eq.	$\theta_t = G\theta_{t-1} + \omega_t$

ν_t, ω_t: mutually independent random variables

Dynamic Linear Model

Vector–Valued State/Observation, Time–Dependent Coefficents

observation eq.	$\boldsymbol{y}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \boldsymbol{\nu}_t$
system eq.	$\boldsymbol{\theta}_t = \boldsymbol{G}_t \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$

$\boldsymbol{\theta}$, $\boldsymbol{\omega}$: n–dimensional random vectors
$\boldsymbol{y}$, $\boldsymbol{\nu}$: r–dimensional random vectors
$\boldsymbol{G}$: $n \times n$ dimensional state evolution matrix
$\boldsymbol{F}$: $n \times r$ dimensional dynamic regression matrix

Dynamic Linear Model

Kalman Filtering

$$\boldsymbol{y}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \boldsymbol{\nu}_t \qquad \boldsymbol{\omega}_t \sim N(\boldsymbol{0}, \boldsymbol{W}_t)$$

$$\boldsymbol{\theta}_t = \boldsymbol{G}_t \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \qquad \boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \boldsymbol{V}_t)$$

$$\theta_0 | D_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$$

post. distr. θ_t: $\theta_t | D_t \sim N(\boldsymbol{m}_t, \boldsymbol{C}_t)$

$$\boldsymbol{m}_t = \boldsymbol{a}_t + \boldsymbol{A}_t \cdot (\boldsymbol{y}_t - \boldsymbol{f}_t)$$
$$\boldsymbol{C}_t = \boldsymbol{R}_t - \boldsymbol{A}_t \cdot \boldsymbol{Q}_t \cdot \boldsymbol{A}_t^T$$
$$\boldsymbol{A}_t = \boldsymbol{R}_t \cdot \boldsymbol{F}_t \cdot \boldsymbol{Q}_t^{-1}$$

forecast: $\boldsymbol{y}_t | D_{t-1} \sim N(\boldsymbol{f}_t, \boldsymbol{Q}_t)$

$$\boldsymbol{f}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{a}_t$$
$$\boldsymbol{Q}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{R}_t \cdot \boldsymbol{F}_t + \boldsymbol{V}_t$$

prior distr. θ_t: $\theta_t | D_{t-1} \sim N(\boldsymbol{a}_t, \boldsymbol{R}_t)$

$$\boldsymbol{a}_t = \boldsymbol{G}_t \cdot \boldsymbol{m}_{t-1}$$
$$\boldsymbol{R}_t = \boldsymbol{G}_t \cdot \boldsymbol{C}_{t-1} \cdot \boldsymbol{G}_t^T + \boldsymbol{W}_t$$

Dynamic Linear Model
Kalman Filtering

$$\boldsymbol{y}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \boldsymbol{\nu}_t$$
$$\boldsymbol{\theta}_t = \boldsymbol{G}_t \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

$$\boldsymbol{\omega}_t \sim N(\boldsymbol{0}, \boldsymbol{W}_t)$$
$$\boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \boldsymbol{V}_t)$$
$$\boldsymbol{\theta}_0 | D_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$$

post. distr. $\boldsymbol{\theta}_t$: $\boldsymbol{\theta}_t | D_t \sim N(\boldsymbol{m}_t, \boldsymbol{C}_t)$

$$\boldsymbol{m}_t = \boldsymbol{a}_t + \boldsymbol{A}_t \cdot (\boldsymbol{y}_t - \boldsymbol{f}_t)$$
$$\boldsymbol{C}_t = \boldsymbol{R}_t - \boldsymbol{A}_t \cdot \boldsymbol{Q}_t \cdot \boldsymbol{A}_t^T$$
$$\boldsymbol{A}_t = \boldsymbol{R}_t \cdot \boldsymbol{F}_t \cdot \boldsymbol{Q}_t^{-1}$$

forecast: $\boldsymbol{y}_t | D_{t-1} \sim N(\boldsymbol{f}_t, \boldsymbol{Q}_t)$

$$\boldsymbol{f}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{a}_t$$
$$\boldsymbol{Q}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{R}_t \cdot \boldsymbol{F}_t + \boldsymbol{V}_t$$

prior distr. $\boldsymbol{\theta}_t$: $\boldsymbol{\theta}_t | D_{t-1} \sim N(\boldsymbol{a}_t, \boldsymbol{R}_t)$

$$\boldsymbol{a}_t = \boldsymbol{G}_t \cdot \boldsymbol{m}_{t-1}$$
$$\boldsymbol{R}_t = \boldsymbol{G}_t \cdot \boldsymbol{C}_{t-1} \cdot \boldsymbol{G}_t^T + \boldsymbol{W}_t$$

Dynamic Linear Model

Kalman Filtering

$$y_t = F_t^T \cdot \theta_t + \nu_t$$
$$\theta_t = G_t \cdot \theta_{t-1} + \omega_t$$

$$\omega_t \sim N(\mathbf{0}, W_t)$$
$$\nu_t \sim N(\mathbf{0}, V_t)$$
$$\theta_0 | D_0 \sim N(m_0, C_0)$$

post. distr. θ_t: $\theta_t | D_t \sim N(m_t, C_t)$

$$m_t = a_t + A_t \cdot (y_t - f_t)$$
$$C_t = R_t - A_t \cdot Q_t \cdot A_t^T$$
$$A_t = R_t \cdot F_t \cdot Q_t^{-1}$$

forecast: $y_t | D_{t-1} \sim N(f_t, Q_t)$

$$f_t = F_t^T \cdot a_t$$
$$Q_t = F_t^T \cdot R_t \cdot F_t + V_t$$

prior distr. θ_t: $\theta_t | D_{t-1} \sim N(a_t, R_t)$

$$a_t = G_t \cdot m_{t-1}$$
$$R_t = G_t \cdot C_{t-1} \cdot G_t^T + W_t$$

Dynamic Linear Model

Kalman Filtering

$$\boldsymbol{y}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \boldsymbol{\nu}_t$$

$$\boldsymbol{\theta}_t = \boldsymbol{G}_t \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

$$\boldsymbol{\omega}_t \sim N(\boldsymbol{0}, \boldsymbol{W}_t)$$

$$\boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \boldsymbol{V}_t)$$

$$\boldsymbol{\theta}_0 | D_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$$

post. distr. $\boldsymbol{\theta}_t$: $\boldsymbol{\theta}_t | D_t \sim N(\boldsymbol{m}_t, \boldsymbol{C}_t)$

$$\boldsymbol{m}_t = \boldsymbol{a}_t + \boldsymbol{A}_t \cdot (\boldsymbol{y}_t - \boldsymbol{f}_t)$$

$$\boldsymbol{C}_t = \boldsymbol{R}_t - \boldsymbol{A}_t \cdot \boldsymbol{Q}_t \cdot \boldsymbol{A}_t^T$$

$$\boldsymbol{A}_t = \boldsymbol{R}_t \cdot \boldsymbol{F}_t \cdot \boldsymbol{Q}_t^{-1}$$

forecast: $\boldsymbol{y}_t | D_{t-1} \sim N(\boldsymbol{f}_t, \boldsymbol{Q}_t)$

$$\boldsymbol{f}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{a}_t$$

$$\boldsymbol{Q}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{R}_t \cdot \boldsymbol{F}_t + \boldsymbol{V}_t$$

prior distr. $\boldsymbol{\theta}_t$: $\boldsymbol{\theta}_t | D_{t-1} \sim N(\boldsymbol{a}_t, \boldsymbol{R}_t)$

$$\boldsymbol{a}_t = \boldsymbol{G}_t \cdot \boldsymbol{m}_{t-1}$$

$$\boldsymbol{R}_t = \boldsymbol{G}_t \cdot \boldsymbol{C}_{t-1} \cdot \boldsymbol{G}_t^T + \boldsymbol{W}_t$$

Dynamic Linear Model
Kalman Filtering

$$\boldsymbol{y}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \boldsymbol{\nu}_t \qquad \boldsymbol{\omega}_t \sim N(\boldsymbol{0}, \boldsymbol{W}_t)$$
$$\boldsymbol{\theta}_t = \boldsymbol{G}_t \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \qquad \boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \boldsymbol{V}_t)$$
$$\boldsymbol{\theta}_0 | D_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0)$$

post. distr. $\boldsymbol{\theta}_t$: $\boldsymbol{\theta}_t | D_t \sim N(\boldsymbol{m}_t, \boldsymbol{C}_t)$

$$\boldsymbol{m}_t = \boldsymbol{a}_t + \boldsymbol{A}_t \cdot (\boldsymbol{y}_t - \boldsymbol{f}_t)$$
$$\boldsymbol{C}_t = \boldsymbol{R}_t - \boldsymbol{A}_t \cdot \boldsymbol{Q}_t \cdot \boldsymbol{A}_t^T$$
$$\boldsymbol{A}_t = \boldsymbol{R}_t \cdot \boldsymbol{F}_t \cdot \boldsymbol{Q}_t^{-1}$$

Bayes' Law

forecast: $\boldsymbol{y}_t | D_{t-1} \sim N(\boldsymbol{f}_t, \boldsymbol{Q}_t)$

$$\boldsymbol{f}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{a}_t$$
$$\boldsymbol{Q}_t = \boldsymbol{F}_t^T \cdot \boldsymbol{R}_t \cdot \boldsymbol{F}_t + \boldsymbol{V}_t$$

prior distr. $\boldsymbol{\theta}_t$: $\boldsymbol{\theta}_t | D_{t-1} \sim N(\boldsymbol{a}_t, \boldsymbol{R}_t)$

$$\boldsymbol{a}_t = \boldsymbol{G}_t \cdot \boldsymbol{m}_{t-1}$$
$$\boldsymbol{R}_t = \boldsymbol{G}_t \cdot \boldsymbol{C}_{t-1} \cdot \boldsymbol{G}_t^T + \boldsymbol{W}_t$$

Glossary

Filtering: Estimate of the current value of the state/system variable.

Smoothing: Estimate of past values of the state/system variable, i.e., estimating at time t given measurements up to time $t' > t$.

Forecasting: Forecasting future observations or values of the state/system variable.

Some Applications

- Modeling of short time series (Bayes model).
- Models composed of different components (trends, seasonality, ARMA)
- Non–stationary models
- Models with interventions
- Regression model with time–dependent coefficients

Some Applications

Regression Model

Regression model: $$y_t = \alpha_t + \beta_{1,t} x_{1,t} + \beta_{2,t} x_{2,t} + \varepsilon_t$$

Some Applications

Regression Model

Regression model: $$y_t = \alpha_t + \beta_{1,t} x_{1,t} + \beta_{2,t} x_{2,t} + \varepsilon_t$$

DLM: $$y_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \varepsilon_t$$
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

Some Applications

Regression Model

Regression model: $$y_t = \alpha_t + \beta_{1,t} x_{1,t} + \beta_{2,t} x_{2,t} + \varepsilon_t$$

DLM:
$$\begin{aligned} y_t &= \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \varepsilon_t \\ \boldsymbol{\theta}_t &= \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \end{aligned}$$

with:
$$\begin{aligned} \boldsymbol{F}_t &= (1, x_{1,t}, x_{2,t})^T \\ \boldsymbol{\theta}_t &= (\alpha_t, \beta_{1,t}, \beta_{2,t})^T \end{aligned}$$

Some Applications

- Modeling of short time series (Bayes model).
- Models composed of different components (trends, seasonality, ARMA)
- Non–stationary models
- Models with interventions
- Regression model with time–dependent coefficients
- ARMA models with known or unknown coefficients

Some Applications

ARMA Model With Known Coefficients

ARMA(2,2) model: $$y_t = \varepsilon_t + \sum_{i=1}^{2} \phi_i y_{t-i} + \sum_{j=1}^{2} \psi_j \varepsilon_{t-j}$$

Some Applications

ARMA Model With Known Coefficients

ARMA(2,2) model:
$$y_t = \varepsilon_t + \sum_{i=1}^{2} \phi_i y_{t-i} + \sum_{j=1}^{2} \psi_j \varepsilon_{t-j}$$

DLM:
$$y_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t$$
$$\boldsymbol{\theta}_t = \boldsymbol{G} \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

Some Applications
ARMA Model With Known Coefficients

ARMA(2,2) model:

$$y_t = \varepsilon_t + \sum_{i=1}^{2} \phi_i y_{t-i} + \sum_{j=1}^{2} \psi_j \varepsilon_{t-j}$$

DLM:

$$y_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t$$

$$\boldsymbol{\theta}_t = \boldsymbol{G} \cdot \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

with:

$$\boldsymbol{F}_t = (1, 0, 0)^T$$

$$\boldsymbol{G} = \begin{pmatrix} \phi_1 & 1 & 0 \\ \phi_2 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\omega}_t = (1, \psi_1, \psi_2)^T \varepsilon_t$$

Some Applications

ARMA Model With Unknown Coefficients

ARMA(2,0) model:

$$y_t = \varepsilon_t + \sum_{i=1}^{2} \phi_i y_{t-i}$$

Some Applications

ARMA Model With Unknown Coefficients

ARMA(2,0) model:

$$y_t = \varepsilon_t + \sum_{i=1}^{2} \phi_i y_{t-i}$$

DLM:

$$y_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \varepsilon_t$$
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

Some Applications

ARMA Model With Unknown Coefficients

ARMA(2,0) model:

$$y_t = \varepsilon_t + \sum_{i=1}^{2} \phi_i y_{t-i}$$

DLM:

$$y_t = \boldsymbol{F}_t^T \cdot \boldsymbol{\theta}_t + \varepsilon_t$$
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

with:

$$\boldsymbol{F}_t = (y_{t-1}, y_{t-2})^T$$
$$\boldsymbol{\theta}_t = (\phi_1, \phi_2)^T$$

Experience

Pros:

- Applicable to short time series
- Not restricted to stationary data
- Large DLM models can be build through composing small DLM models
- May be extended to non-normal distributions
- Results are easy to interpret

Cons:

- Finding size of noise terms ν_t, ω_t difficult
  - MLE yields unreasonable results frequently.
  - A solution might be Bayesian estimation of ν_t, ω_t which is not part of the R–libraries
- Incomplete R libraries, no “standard” library
- Difficult numerical implementation

R–Libraries

- **dlm** Personally preferred library, see Petris (2010)
- **FKF** Performance–optimized Kalman filtering, no functions for model building
- **sspir** Library used by Cowpertwait u. Metcalfe (2009), no longer maintained
- **dlmodeler** Common interface to libraries dlm, FKF, KFAS
- **. . .** Several other libraries, see CRAN Task View: "Time Series Analysis", for comparison see Tusell (2011) and Commandeur u. a. (2011)..

Some References I

[Commandeur u. a. 2011] Commandeur, Jacques J. F. ; Koopman, Siem J. ; Ooms, Marius: Statistical Software for State Space Methods. In: Journal of Statistical Software 41 (2011), 5, Nr. 1, 1–18. http://www.jstatsoft.org/v41/i01. – ISSN 1548–7660

[Cowpertwait u. Metcalfe 2009] Cowpertwait, Paul S. P. ; Metcalfe, Andrew V.: Introductory Time Series with R. Springer, 2009 (Use R!). – 256 S. http://dx.doi.org/10.1007/978-0-387-88698-5. http://dx.doi.org/10.1007/978-0-387-88698-5. – ISBN 978–0–387–88697–8

[Dethlefsen u. Lundbye-Christensen 2006] Dethlefsen, Claus ; Lundbye-Christensen, Søren: Formulating State Space Models in R with Focus on Longitudinal Regression Models. In: Journal of Statistical Software 16 (2006), Mai, Nr. 1, 1–15. http://www.jstatsoft.org/v16/i01/paper, Abruf: 2011-07-25

[Durbin u. Koopman 2002] Durbin, J. ; Koopman, S.J.: Time Series Analysis by State Space Methods: Second Edition. OUP Oxford, 2002 (Oxford Statistical Science Series). http://books.google.de/books?id=fOq39Zh0olQC. – ISBN 9780199641178

Some References II

[Honerkamp 1990] HONERKAMP, J.: Stochastische Dynamische Systeme. Weinheim : vch, 1990

[Hyndman 2009] HYNDMAN, Rob J.: forecast: Forecasting functions for time series. 1.25, 2009. `http://CRAN.R-project.org/package=forecast`

[Petris 2010] PETRIS, Giovanni: An R Package for Dynamic Linear Models. In: Journal of Statistical Software 36 (2010), Nr. 12, 1–16. `http://www.jstatsoft.org/v36/i12/paper`, Abruf: 2010-10-15

[Petris u. Petrone 2011] PETRIS, Giovanni ; PETRONE, Sonia: State Space Models in R. In: Journal of Statistical Software 41 (2011), Mai, Nr. 4, 1–25. `http://www.jstatsoft.org/v41/i04/paper`, Abruf: 2011-05-20

[Petris u. a. 2007] PETRIS, Giovanni ; PETRONE, Sonia ; CAMPAGNOLI, Patrizia: Dynamic Linear Models with R. Version: aug 2007. `http://www.math.u-bordeaux1.fr/~pdelmora/dynamics-linear-models.petris_et_al.pdf`, Abruf: 2013-07-10. – Preprint

[Tusell 2011] TUSELL, Fernando: Kalman Filtering in R. In: Journal of Statistical Software 39 (2011), 3, Nr. 2, 1–27. `http://www.jstatsoft.org/v39/i02`. – ISSN 1548–7660

Some References III

[West u. Harrison 1997] West, Mike ; Harrison, Jeff: Bayesian Forecasting and Dynamic Models. 2. Springer-Verlag, 1997 (Springer Series in Statistics). – 680 S. – ISBN 0–387–94725–6

Questions?

Thank You! ☺

➜ LUX