Robust time series analysis with R

Alexander Dürre

ECARES - European Center for Advanced Research in Economics and Statistics
Université Libre de Bruxelles

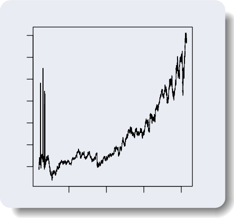
Wien, 20th September 2021

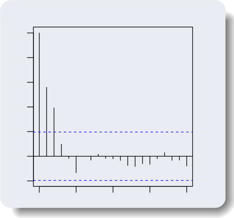
R packages for robust time series analysis

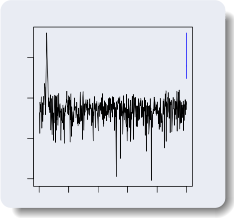
Package	Content
npcp	change-point analysis
robcp	change-point analysis
robets	forcasting
robfilter	time series filter
robustarima	parametric modelling
RobKF	model based filtering
RobPer	periodogram estimation
robts	basic time series analysis

Robust time series packages in R

Basic descriptive statistics







Autocorrelation function

Let $(X_t)_{t\in\mathbb{N}_0}$ be a stationary time series

Theoretical ACF

If $\mathbb{E}(X_0^2) < \infty$ define $\gamma : \mathbb{N}_0 \to [-1, 1]$ by

$$\rho(h) = \mathrm{Cor}(X_0, X_h) = \frac{[\mathbb{E}(X_h) - \mu][\mathbb{E}(X_0) - \mu]}{\mathrm{Var}(X_0)}$$

where $\mu = \mathbb{E}(X_0)$

Empirical ACF

$$\hat{\rho}(h) = \frac{\sum_{i=1}^{n-h}[X_i - \overline{X}][X_{i+h} - \overline{X}]}{\sum_{i=1}^{n}[X_i - \overline{X}]^2}$$

Autocorrelation function

Let $(X_t)_{t\in\mathbb{N}_0}$ be a stationary time series

Definition ACF

If $\mathbb{E}(X_0^2) < \infty$ define $\rho : \mathbb{N}_0 \to [-1, 1]$ by

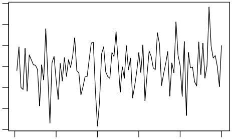
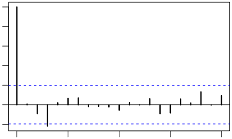
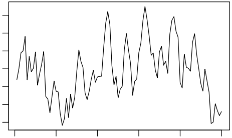
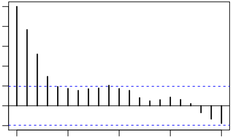
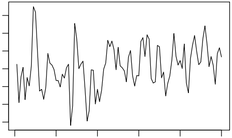
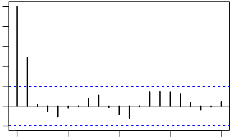
$$\rho(h) = \text{Cor}(X_0, X_h) = \frac{[\mathbb{E}(X_h) - \mu][\mathbb{E}(X_0) - \mu]}{\text{Var}(X_0)}$$

where $\mu = \mathbb{E}(X_0)$

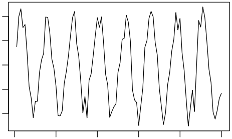
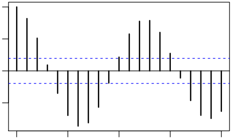
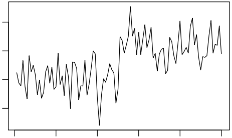
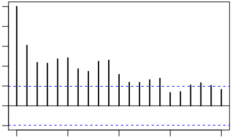
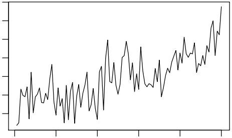
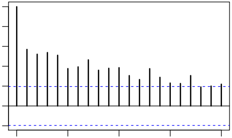
stats

```
> acf(x)
```

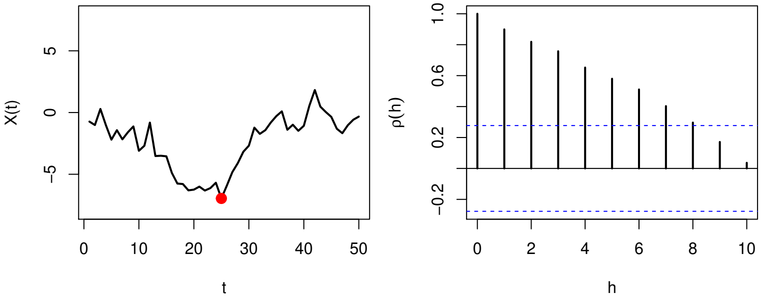
Typical ACF shapes

Process	Line chart	ACF
White noise		
Autoregressive process		
Moving average process		

Typical ACF shapes II

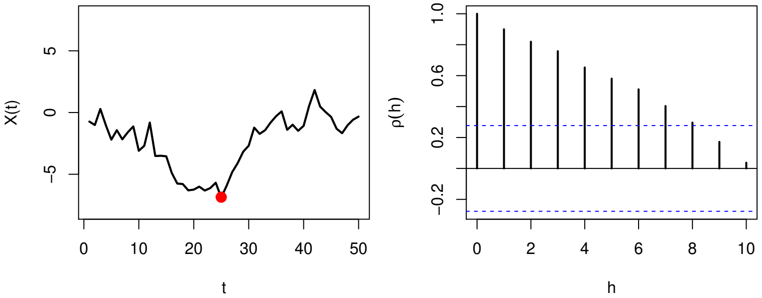
Process	Line chart	acf
Periodicity		
Change point		
Trend		

Sensitivity of ACF



AR(1) with $\pi = 0.9$ and one outlier (left) and estimated ACF (right)

Sensitivity of ACF



AR(1) with $\pi = 0.9$ and one outlier (left) and estimated ACF (right)

Robust alternatives to ACF

- based on variances (Ma and Genton, 2000)

robts

```
> acfrob(x,approach="GK")
```

- based on robustly filtered timeseries (Maronna et al., 2006)

robts

```
> acfrob(x,approach="filter")
```

- based on multivariate correlation matrices (Dürre et al., 2015)

robts

```
> acfrob(x,approach="multi")
```

- based on sign and ranks (Mottonen et al., 1999)

robts

```
> acfrob(x,approach="rank")
```

ACF based on variances (Ma and Genton, 2000)

Based on the identity

$$\rho(h) = \frac{\text{Var}(X_1 + X_{1+h}) - \text{Var}(X_1 - X_{1+h})}{\text{Var}(X_1 + X_{1+h}) + \text{Var}(X_1 - X_{1+h})}$$

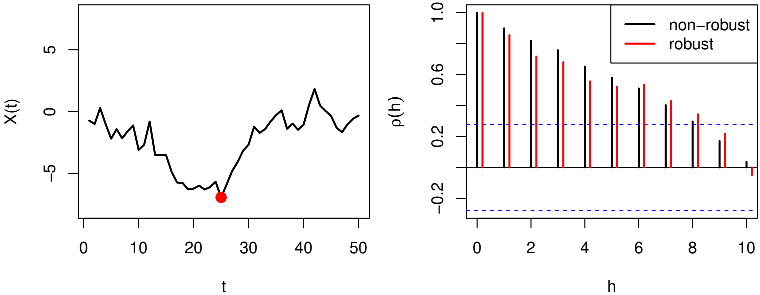
Now you can use any (robust) variance estimator!

robts

```
> acfrob(x)
```

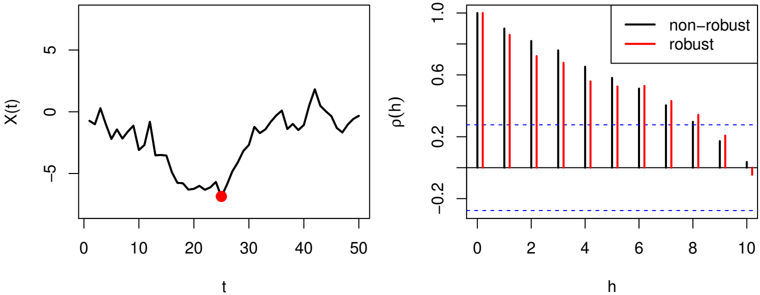
uses variance approach with *Qn* since its is fast and robust in most circumstances

Sensitivity of ACF II



AR(1) with parameter $\pi = 0.9$ and one outlier (left) and estimated ACF [non-robust - black; robust - red] (right)

Sensitivity of ACF II



AR(1) with parameter $\pi = 0.9$ and one outlier (left) and estimated ACF [non-robust - black; robust - red] (right)

Spectrum

Let $\mathbb{E}(X_0^2) < \infty$ and $\sum_{i=1}^{\infty} |\rho(h)| < \infty$, define $S : [0, \frac{1}{2}] \to \mathbb{R}_+$

Spectrum

$$S(f) = \text{Var}(X_0) \sum_{-\infty}^{\infty} \rho(k) e^{-2\pi i k f} = \text{Var}(X_0) \left[1 + \sum_{k=1}^{\infty} \cos(2\pi f k) \rho(k) \right]$$

Estimation

$$\hat{S}(f) = \frac{1}{T} |Z(f)|^2 \text{ with } Z(f) = \sum_{i=1}^{T} \tilde{X}_t e^{-2\pi i f t}$$

where $\tilde{X}_t = (X_t - \overline{X}) d(t/n)$ and $d : [0,1] \to [0,1]$ is a taper

Spectrum

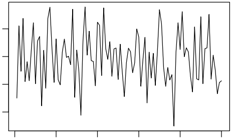
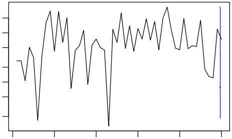
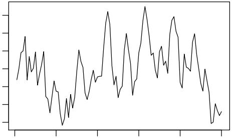
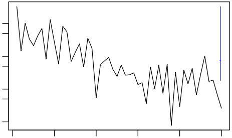
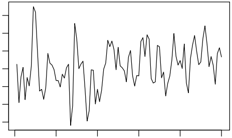
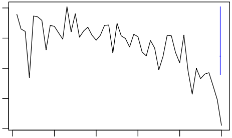
Let $\mathbb{E}(X_0^2) < \infty$ and $\sum_{i=1}^{\infty} |\rho(h)| < \infty$, define $S : [0, \frac{1}{2}] \to \mathbb{R}_+$

Definition

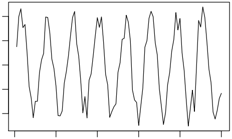
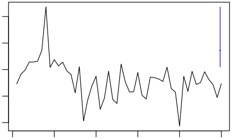
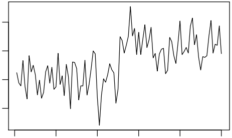
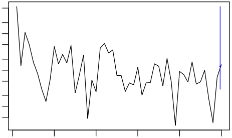
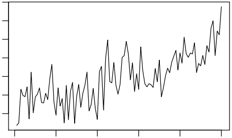
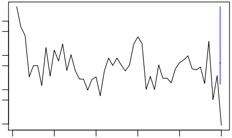
$$S(f) = \mathrm{Var}(X_0) \sum_{-\infty}^{\infty} \rho(k) e^{-2\pi i k f} = \mathrm{Var}(X_0) \left[1 + \sum_{k=1}^{\infty} \cos(2\pi f k) \rho(k) \right]$$

stats

```
> spectrum(x)
```

Process	Line chart	Spectrum
White noise		
Autoregressive process		
Moving average process		

Typical spectrum shapes II

Process	Line chart	Spectrum
Periodicity		
Change point		
Trend		

Robust alternatives to empirical spectrum

- AR fit + Periodogram of residuals (Maronna et al., 2006)

robts

```
> spectrumrob(x,method="pgram")
```

- Fouriertransform of robust acf (Spangl, 2008)

robts

```
> spectrumrob(x,method="acf")
```

- Fitting periodic functions by robust regression (Thieler et al., 2013)

RobPer

```
> n <- length(x)
> freq <- seq(from=1/n,to=0.5,by=1/n)
> RobPer(cbind(1:n,x),weighting=FALSE,periods
  =1/freq,regression="S", model="sine")
```

Robust periodogram based on AR fit

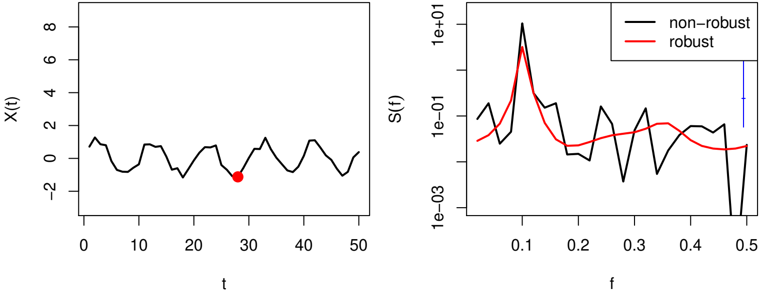
robts

```
> spectrumrob(x)
```

gives smoothed periodogram! (alternatively: set bandwidth=0)

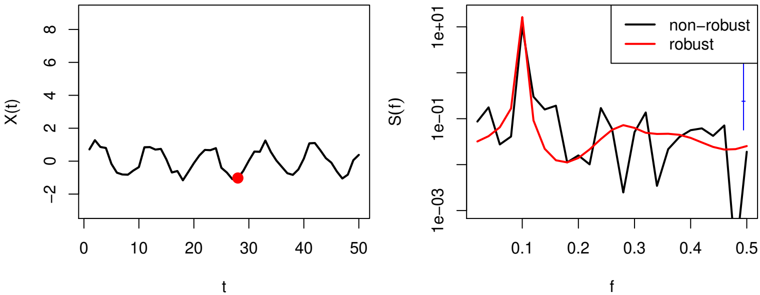
1. Fit an AR model
2. Compute periodogram of parametric model $\hat{S}_{AR}(f)$
3. Calculate residuals $(\hat{\epsilon}_i)_{i=1,\ldots,n}$
4. Robustify residuals $(\tilde{\epsilon}_i)_{i=1,\ldots,n}$
5. Compute periodogram of $(\tilde{\epsilon}_i)_{i=1,\ldots,n}$: $\hat{S}_\epsilon(f)$
6. Combine both estimations: $\hat{S} = \hat{S}_{AR}(f) \cdot \hat{S}_\epsilon(f)$

Sensitivity of the Periodogram



Time series consisting of a sine, white noise and one outlier (left) and estimated spectrum [non-robust - black; robust - red] (right)

Sensitivity of the Periodogram



Time series consisting of a sine, white noise and one outlier (left) and estimated spectrum [non-robust - black; robust - red] (right)

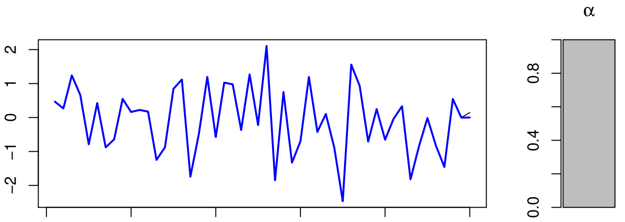
Smoothing

- Aim: separate
  signal (low freq, comp.) from **noise** (high freq. comp.)
- Easiest scenario: $X_t = \mu + \epsilon_t$ where $(\epsilon_t)_{t\in\mathbb{N}}$ is iid
  $\Rightarrow s_t = \mu$ is signal $n_t = \epsilon_t$ is noise

Simple exponential smoothing

For a smoothing parameter $\alpha \in (0,1)$:

$$\hat{s}_t = \hat{s}_{t-1} + \alpha(X_t - \hat{s}_{t-1}) = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i X_{t-i}$$



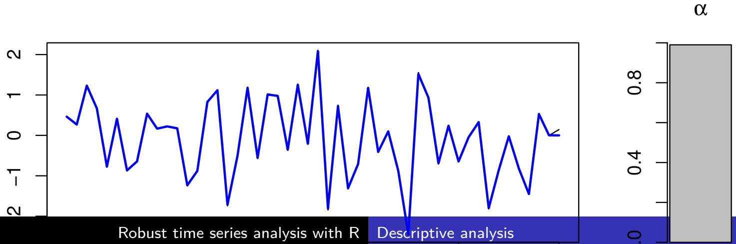
Smoothing

- Aim: separate
  signal (low freq, comp.) from **noise** (high freq. comp.)
- Easiest scenario: $X_t = \mu + \epsilon_t$ where $(\epsilon_t)_{t \in \mathbb{N}}$ is iid
  $\Rightarrow s_t = \mu$ is signal $n_t = \epsilon_t$ is noise

Simple exponential smoothing

For a smoothing parameter $\alpha \in (0, 1)$:

$$\hat{s}_t = \hat{s}_{t-1} + \alpha(X_t - \hat{s}_{t-1}) = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i X_{t-i}$$



More complex exponential smoothing

Simple exponential smoothing:

$$\hat{s}_t = \hat{s}_{t-1} + \alpha \underbrace{(X_t - \hat{s}_{t-1})}_{\hat{\epsilon}_t}$$

corresponds to the state-space model:

$$X_t = s_{t-1} + \epsilon_t \text{ with } s_t = s_{t-1} + \alpha\epsilon_t$$

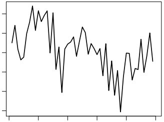
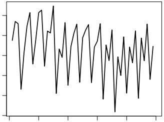
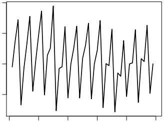
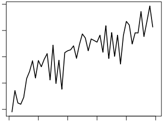
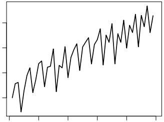
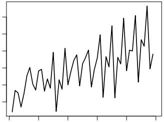
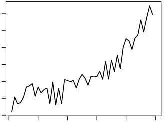
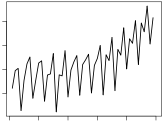
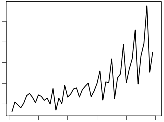
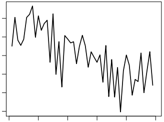
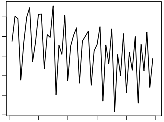
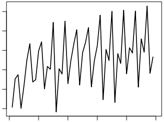
Estimate α with (conditional) likelihood assuming $\epsilon_t \sim N(0, \sigma^2)$.

Generalized state space models

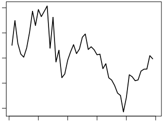
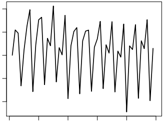
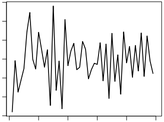
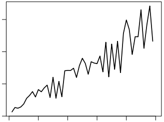
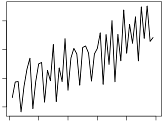
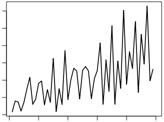
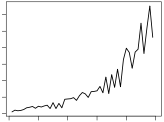
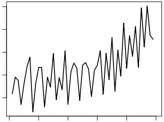
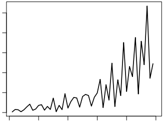
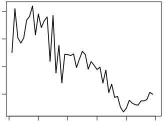
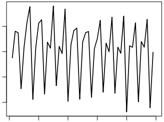
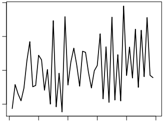
$$\begin{aligned} X_t &= h(s_t) + k(s_t)\epsilon_t \\ s_t &= f(s_{t-1}) + d(s_{t-1})\epsilon_t \end{aligned}$$

where $(\epsilon_t)_{t\in\mathbb{Z}}$ is iid with $\epsilon_0 \sim N(0, 1)$

More complex **A**ddative error models

Trend \ Season	**N**one	**A**dditive	**M**ultiplicative
None			
Additive			
Multiplicatice			
Damped			

More complex Multiplicative error models

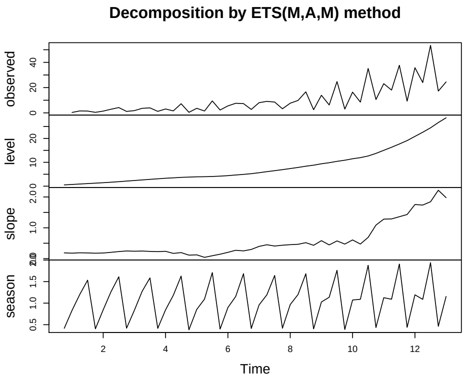
Trend \ Season	**N**one	**A**dditive	**M**ultiplicative
None			
Additive			
Multiplicatice			
Damped			

Estimation of smoothing parameter / state space model

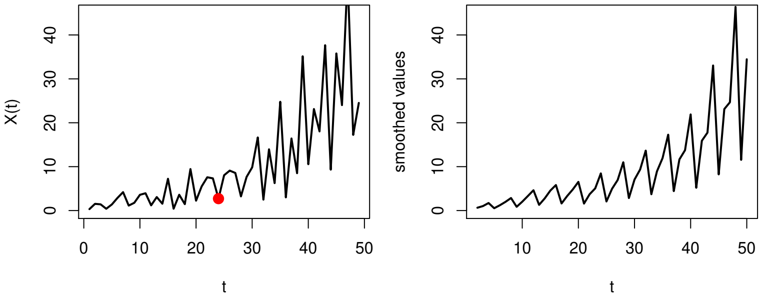
forecast

```
> ets(x)
```

- chooses best model by (corrected) AIC criterion
- periodicity is chosen by `freq` specified in `ts` object!

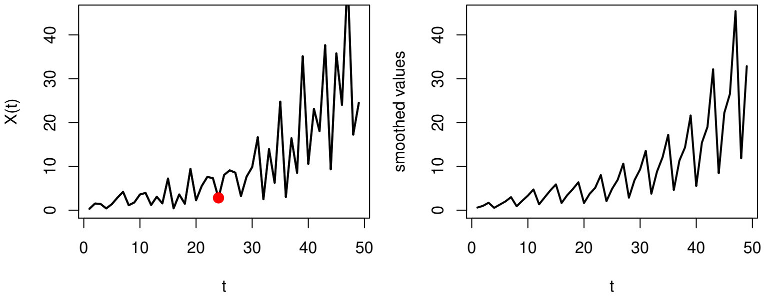


Sensitivity of exponential smoothing



Time series from an MMM-model with one outlier (left) and estimated signal using ets (right)

Sensitivity of exponential smoothing



Time series from an MMM-model with one outlier (left) and estimated signal using ets (right)

Robust alternatives to smoothing

- Filter based on repeated median for models with trend and jumps (Fried, 2004)

robfilter

```
> robust.filter(x,width=bandwidth)
```

- Robust version of ets (Crevits and Croux, 2017)

robets

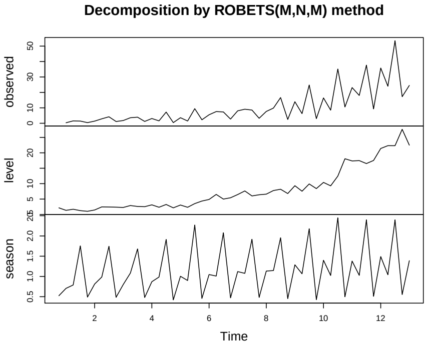
```
> robets(x)
```

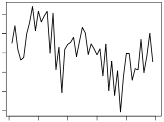
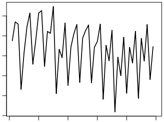
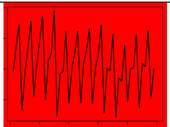
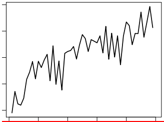
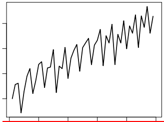
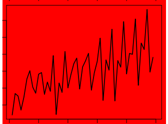
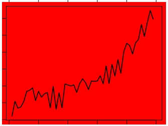
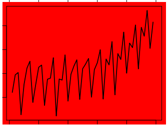
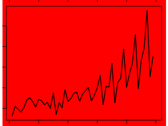
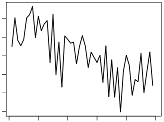
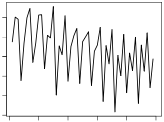
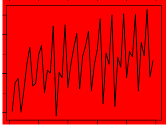
Package robets

robets

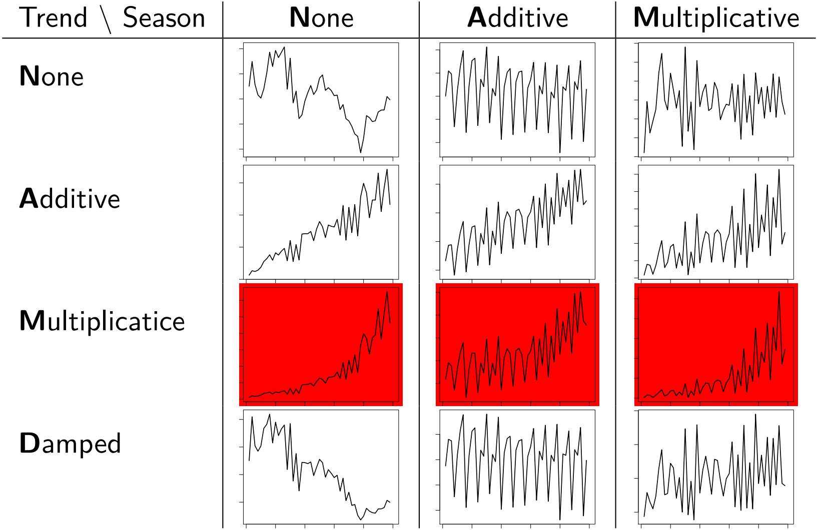
```
> robets(x)
```

1. Outlier cleaning based on state space model
2. Parameter estimation by optimization of robustified likelihood: Robust scale instead of sum of squared residuals
3. Model selection by robustified AIC criterion

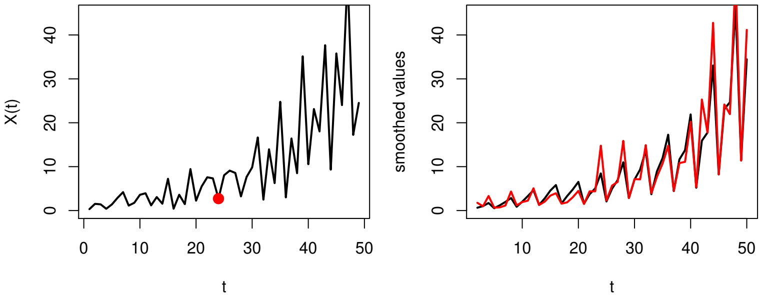


Trend \ Season	None	Additive	Multiplicative
None			
Additive			
Multiplicatice			
Damped			

Implemented Multiplicative error models

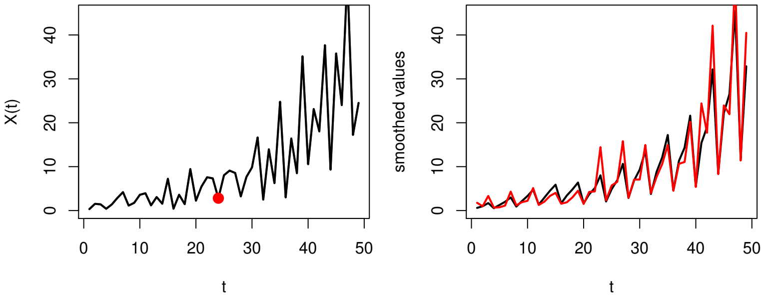


Sensitivity of smoothing



Time series from an MMM-model with one outlier (left) and estimated signal using ets [black] respectively robets [red] (right)

Sensitivity of smoothing



Time series from an MMM-model with one outlier (left) and estimated signal using ets [black] respectively robets [red] (right)

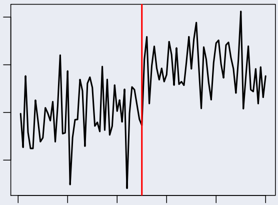
Exercises A

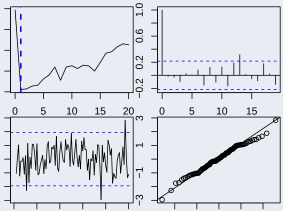
1. Calculate the ACF of the stock prices of SAP (SAP2.Rdata) robustly and by the empirical ACF. consider the raw prices and the log returns defined by $(\tilde{X}_t)_{t=2,\dots,n}$ defined by $\tilde{X}_t = \log(X_t) - \log(X_{t-1})$).
2. Determine the frequency of the guitar string (guitar2.Rdata)! (Which string is played and is it tuned correctly?)
3. Smooth the NO2 values (NO2Krefeld.Rdata) robustly and non robustly. Decide for a reasonable period length (and specify it by the freq argument in the ts-object)
4. *additional:* Investigate the influence of block outliers in the estimation of the ACF. Start with a time series of iid random variables and add an increasing value to 5 consecutive observations.

Note: install `robts` (after installing dependencies: robustbase, rrcov, SpatialNP, ICSNP, sscor, quantreg, ltsa) by:

```
install.packages("robts", repos="http://R-Forge.R-project.org")
```

Modelling





Searching for changes

Assume: $(X_i)_{i\in\mathbb{N}}$

Test for change in mean

$$H_0: \ \mathbb{E}(X_1) = \ldots = \mathbb{E}(X_n)$$

against

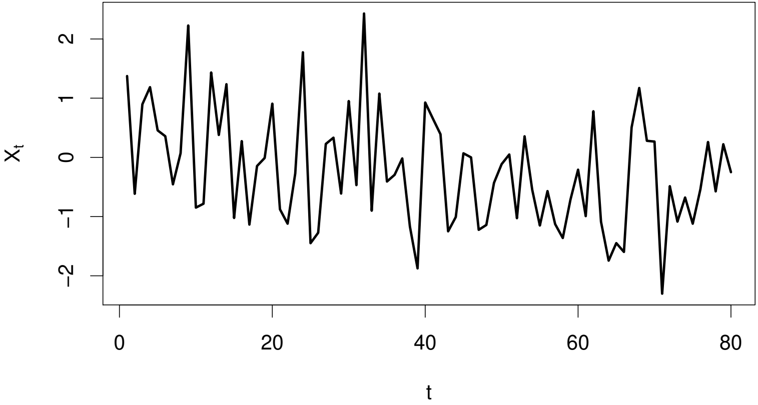
$$H_1: \ \exists k: \ \mathbb{E}(X_{k-1}) \neq \mathbb{E}(X_k)$$

Estimate also time of change

Example: change in location

Simulated example:

$$X_t \sim \begin{cases} N(0,1) & 1 \leq t \leq 40 \\ N(-1,1) & 41 \leq t \leq 80 \end{cases}$$

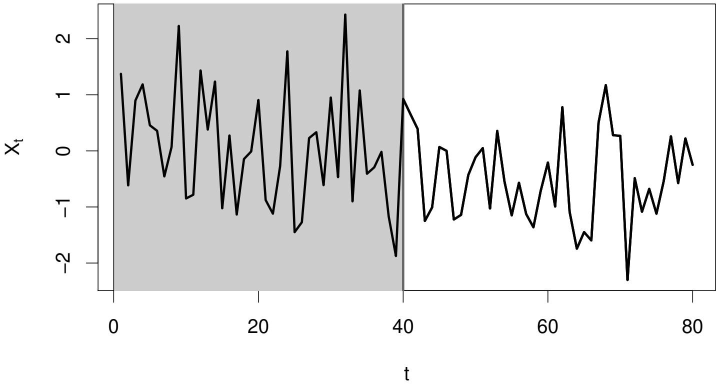


Time series with jump in $k = 41$

Time of change known

2 sample t-test:

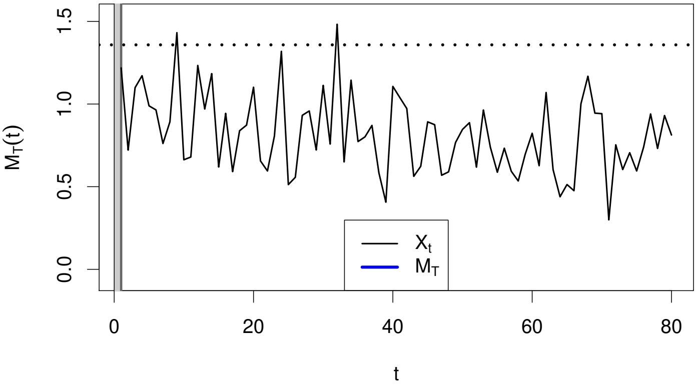
$$\sqrt{\frac{k(T-k)}{T}}\left(\frac{\overline{X}_{[1:40]}-\overline{X}_{[41:80]}}{\hat{\sigma}}\right)\sim N(0,1)$$



Time of change unknown

Series of two sample tests:

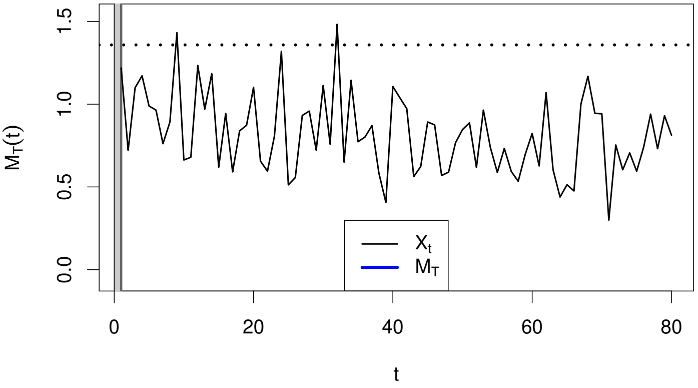
$$M_T(t) = \frac{t(T-t)}{T^{3/2}\hat{\sigma}}(\overline{X}_{[1:t]} - \overline{X}_{[(t+1:T)]}) = \frac{1}{\sqrt{T}\hat{\sigma}}\left(\sum_{k=1}^{t} X_k - \frac{t}{T}\sum_{k=1}^{T} X_k\right)$$



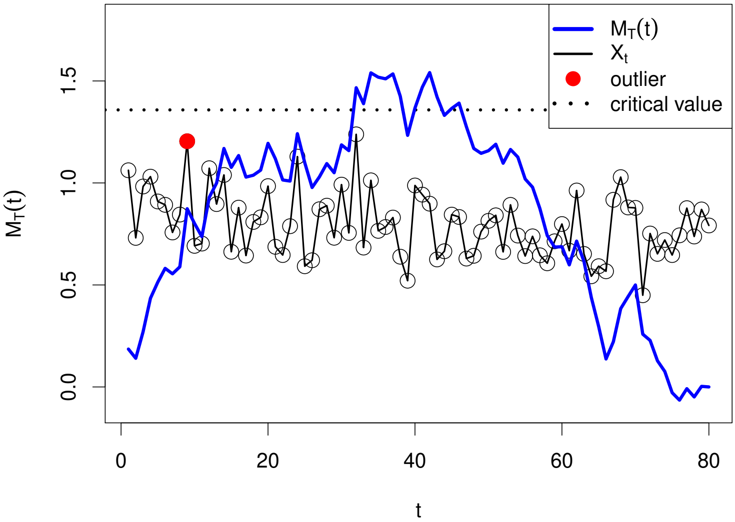
Time of change unknown

Series of two sample tests:

$$M_T(t) = \frac{t(T-t)}{T^{3/2}\hat{\sigma}}(\overline{X}_{[1:t]} - \overline{X}_{[(t+1:T)]}) = \frac{1}{\sqrt{T}\hat{\sigma}}\left(\sum_{k=1}^{t} X_k - \frac{t}{T}\sum_{k=1}^{T} X_k\right)$$

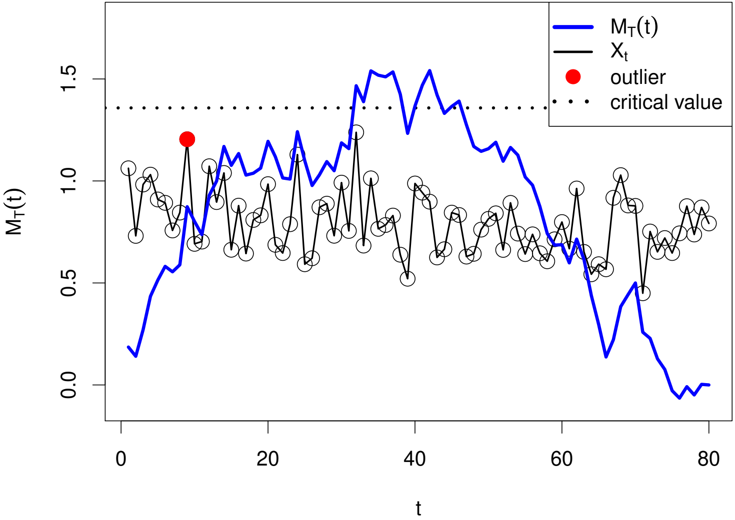


Cusum statistic with one outlier



Cusum trajectory in case of one outlier

Cusum statistic with one outlier

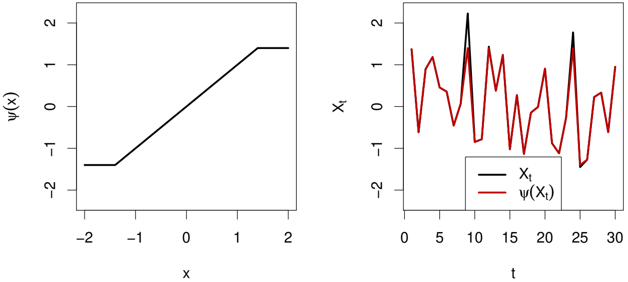


Cusum trajectory in case of one outlier

Robustification

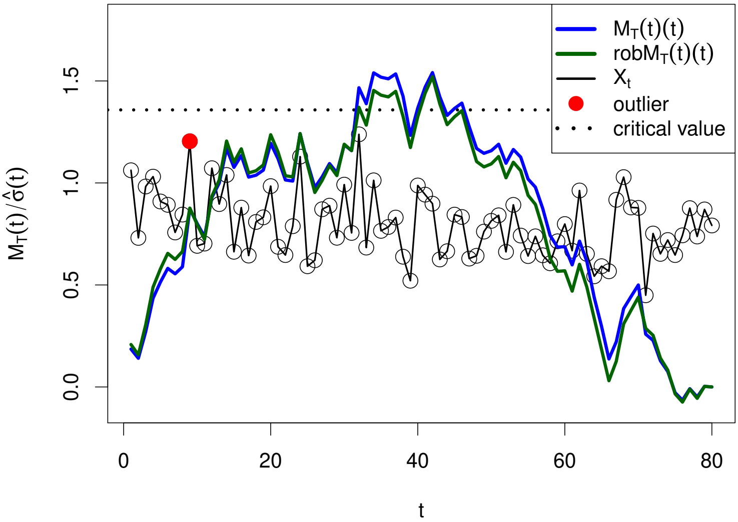
Intuitive solution to bound the influence of one observation:
Transform values by a bounded function e.g:

$$\psi_H(x) = \begin{cases} x & |x| \leq k \\ -k & x < -k \\ k & x > k \end{cases}$$



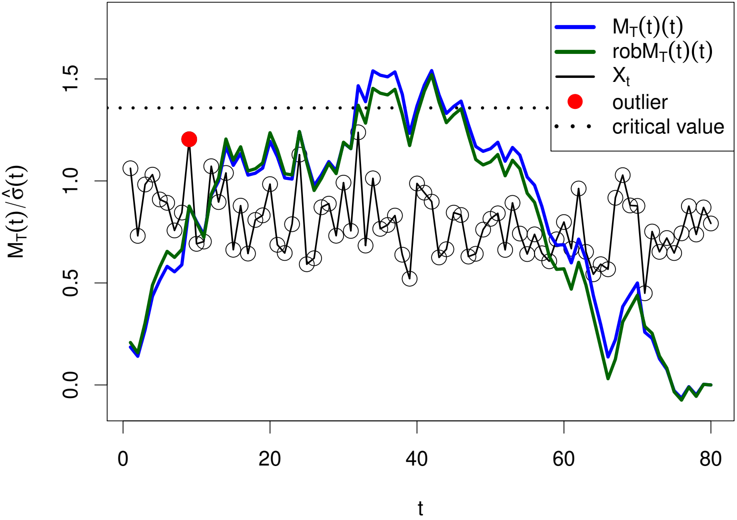
Possible $\psi-$function (left) and transformed time series (right)

Robust cusum statistic with one outlier



Cusum trajectories in case of one outlier

Robust cusum statistic with one outlier



Cusum trajectories in case of one outlier

Resulting robust cusum statistic

$$S_T = \sup_{x \in [0,1]} \frac{1}{\sqrt{T}\hat{\sigma}} \left| \sum_{i=1}^{\lfloor Tx \rfloor} \psi\left(\frac{X_i - \hat{\mu}}{\hat{v}}\right) - \frac{\lfloor Tx \rfloor}{T} \sum_{i=1}^{T} \psi\left(\frac{X_i - \hat{\mu}}{\hat{v}}\right) \right|$$

where

- $\psi : \mathbb{R} \to \mathbb{R}$ is bounded
- $\hat{\mu}$ is a location estimator, e.g. median
- $\hat{v}$ is a scale estimator, e.g. MAD
- $\hat{\sigma}^2$ estimator for long run variance

R-functions

- Non-robust Cusum-test

robcp

```
> cusumstat <- CUSUM(x)
> pKSdist(cusumstat)
```

- Location based on Hodges Lehmann (Dehling et al., 2020) [not to large *n*!]

robcp

```
> hl_test(x)
```

- Huberized test for change in location respectively variance (Dürre and Fried, 2019)

robcp

```
> huber_cusum(x,fun="HLm")
> huber_cusum(x,fun="HCm")
```

Fitting ARMA-models

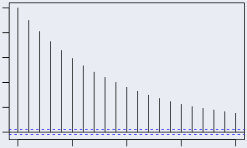
ARMA Model

for $p, q \geq 0$:

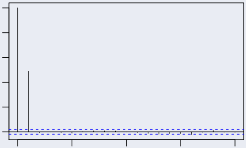
$$X_t = \pi_1 X_{t-1} + \ldots + \pi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

where $(\epsilon_t)_{t \in \mathbb{Z}}$ is a sequence of iid random variables (often $N(0, \sigma^2)$)

acf AR(1)



acf MA(1)



Fitting ARMA-models

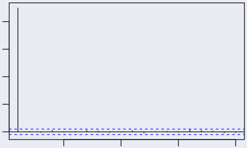
ARMA-model

for $p, q \geq 0$:

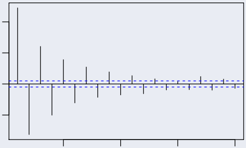
$$X_t = \pi_1 X_{t-1} + \ldots + \pi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

where $(\epsilon_t)_{t \in \mathbb{Z}}$ is a sequence of iid random variables (often $N(0, \sigma^2)$)

pacf AR(1)



pacf MA(1)



Non-robust estimation

3 step procedure of Hannan and Rissanen (1982):

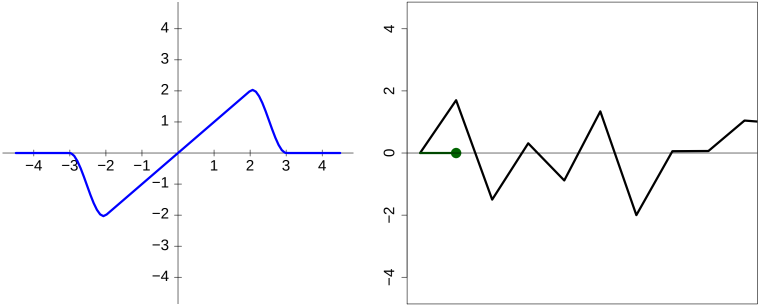
1. Fit AR model of large order p_0 by Yule-Walker equations $(\hat{\rho} \to \hat{\pi})$
2. Estimate parameter by

$$\tilde{\sigma}^2_{p,q} = \inf \frac{1}{n} \sum_{i=p_0+\max(p,q)}^{n} \left(X_i - \sum_{j=1}^{p} \pi_j X_{i-j} - \sum_{j=1}^{q} \theta_j \hat{\epsilon}_{i-j} \right)^2$$

   where $(\hat{\epsilon}_i)_{i=p_0}^{n}$ are estimated from the inital AR fit
3. Choose p, q by minimizing an aic criterion based on $\tilde{\sigma}_{p,q}$
4. Update parameter estimations with new estimated residuals

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

- **prediction $\hat{X}_t$ based on filtered values $Y_{t-1}, \ldots Y_{t-p}$**
- residual $\epsilon_t = X_t - \hat{X}_t$
- filtered value $Y_t = \hat{X}_t + \psi(\epsilon_t)$

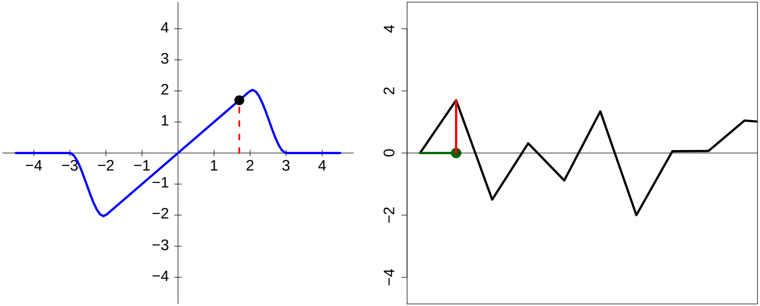


Possible Ψ−function (left) and filtered time series (right)

Robust filter, Masreliez (1975) and Maronna et al. (2006)

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

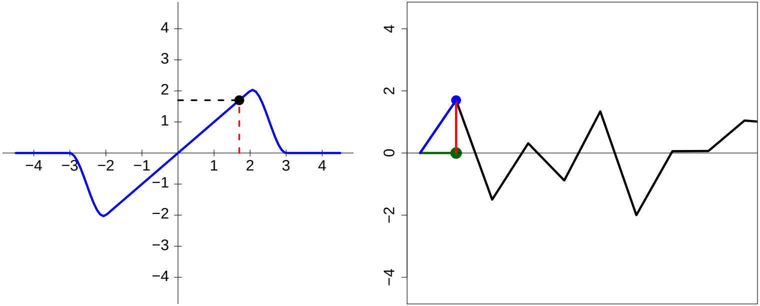
- prediction $\hat{X}_t$ based on filtered values $Y_{t_1}, \ldots Y_{t-p}$
- **residual** $\epsilon_t = X_t - \hat{X}_t$
- filtered value $Y_t = \hat{X}_t + \psi(\epsilon_t)$



Possible Ψ−function (left) and filtered time series (right)

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

- prediction $\hat{X}_t$ based on filtered values $Y_{t-1}, \ldots Y_{t-p}$
- residual $\epsilon_t = X_t - \hat{X}_t$
- **filtered value** $Y_t = \hat{X}_t + \psi(\epsilon_t)$

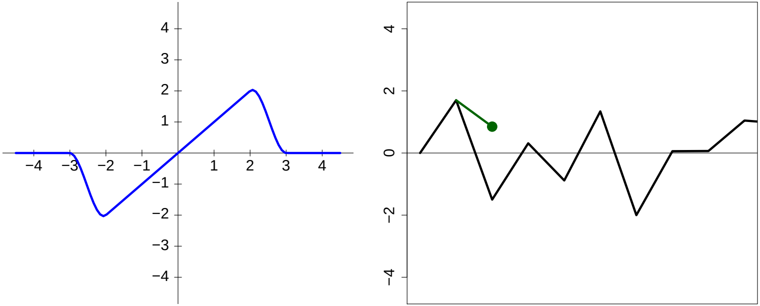


Possible $\Psi-$function (left) and filtered time series (right)

Robust filter, Masreliez (1975) and Maronna et al. (2006)

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

- **prediction $\hat{X}_t$ based on filtered values $Y_{t-1}, \ldots Y_{t-p}$**
- residual $\epsilon_t = X_t - \hat{X}_t$
- filtered value $Y_t = \hat{X}_t + \psi(\epsilon_t)$

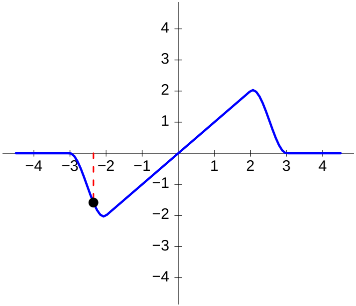


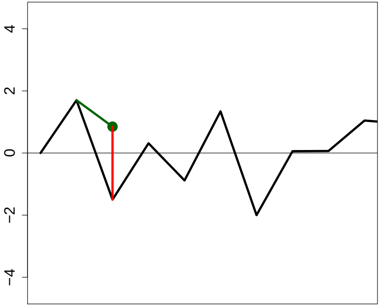
Possible Ψ−function (left) and filtered time series (right)

Robust filter, Masreliez (1975) and Maronna et al. (2006)

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

- prediction $\hat{X}_t$ based on filtered values $Y_{t-1}, \ldots Y_{t-p}$
- **residual** $\epsilon_t = X_t - \hat{X}_t$
- filtered value $Y_t = \hat{X}_t + \psi(\epsilon_t)$

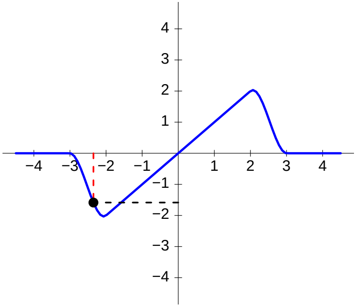


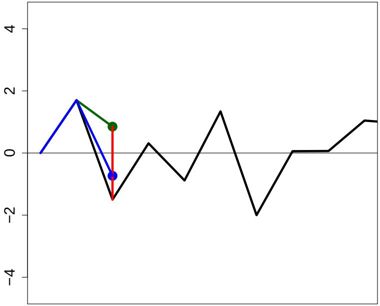


Possible $\Psi-$function (left) and filtered time series (right)

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

- prediction $\hat{X}_t$ based on filtered values $Y_{t-1}, \ldots Y_{t-p}$
- residual $\epsilon_t = X_t - \hat{X}_t$
- **filtered value** $Y_t = \hat{X}_t + \psi(\epsilon_t)$



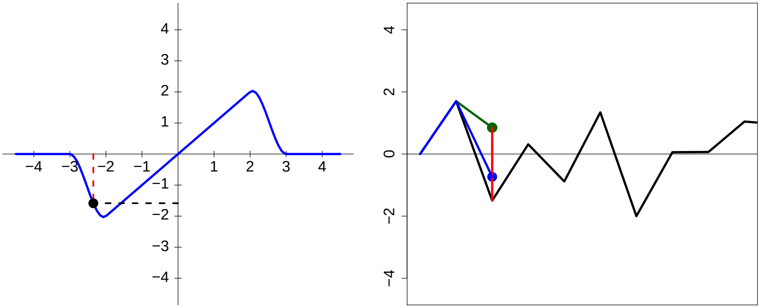


Possible $\Psi-$function (left) and filtered time series (right)

Robust filter, Masreliez (1975) and Maronna et al. (2006)

For given AR model with $\pi_1, \ldots, \pi_p$ calculate iteratively:

- prediction $\hat{X}_t$ based on filtered values $Y_{t-1}, \ldots Y_{t-p}$
- residual $\epsilon_t = X_t - \hat{X}_t$
- filtered value $Y_t = \hat{X}_t + \psi(\epsilon_t)$



Possible $\Psi-$function (left) and filtered time series (right)

Robust filter, Masreliez (1975) and Maronna et al. (2006)

Since AR model is unknown:

1. Minimize robust BIC with respect to AR parameter $\pi_1, \ldots, \pi_{p_0}$ to finit initial *AR* process

$$\log(\hat{\sigma}^2_{\pi_1,\ldots,\pi_{p_0}}(\epsilon_{p_0}, \ldots, \epsilon_n)) + \frac{\log(n)p}{n}$$

2. Estimate π and θ by robustified likelihood

$$\tilde{\sigma}^2_{p,q} = \inf \hat{\sigma}^2(\hat{\epsilon}_{p_0+\max p,q}, \ldots, \hat{\epsilon}_n)$$

   where $\hat{\epsilon}_i = X_i - \sum_{j=1}^{p} \pi_j X_{i-j} - \ldots \sum_{j=1}^{q} \theta_j \epsilon_{i-j}$

3. Choose p, q by robustified aic criterion based on $\tilde{\sigma}^2_{p,q}$.

R-functions

- Classical ARMA estimation

tseries

```
> arma(x)
```

- Robustified estimation of AR models

robts

```
> arrob(x)
```

- Robustified estimation of ARMA models

robts

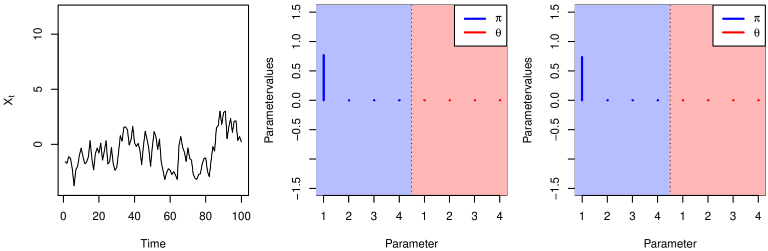
```
> armarob(x,arorder=maxp,maorder=maxq,aic=TRUE,
  aicpenalty=function(p) return(p*log(length(x))))
```

- robustified estimation of ARIMA models

robustarima

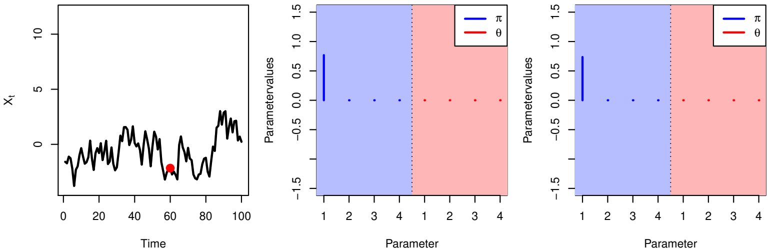
```
> dat <- data.frame(1:n,x)
> rob.arima(dat,x~1,p=p,q=q)
```

Sensitivity of ARMA estimation



AR(1) with $\pi = 0.8$ and one outlier (left), parameters of non robustly estimated ARMA (center) and parameters of robustly estimated ARMA (right). Parameters of AR part in blue and parameters of MA part in red.

Sensitivity of ARMA estimation

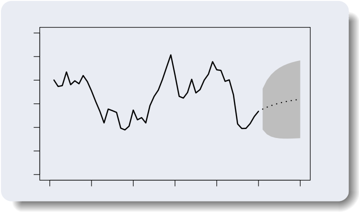


AR(1) with $\pi = 0.8$ and one outlier (left), parameters of non robustly estimated ARMA (center) and parameters of robustly estimated ARMA (right). Parameters of AR part in blue and parameters of MA part in red.

Exercises

1. Detect changes in the location of the log returns of the SAP stock prices. If you can detect a change (significance level 0.05), split the time series in two and try to detect changes there.
2. Fit an ARMA model to the yearly sunhours of Chemnitz robustly, try also conventional fits of order (1,0),(0,1) an (1,1).

Forecasting



Non-robust methods

- Using the state space model of smoothing approach (**?**)

forcast

```
> modelfit <- ets(x)
> predict(modelfit,n.ahead=predtime)
```

- Using the ARMA fit

stats

```
> armafit <- arima(x,c(p,0,q))
> predict(armafit,n.ahead=predtime)
```

Robust methods

- Using the state space model of smoothing approach (Crevits and Croux, 2017)

robets

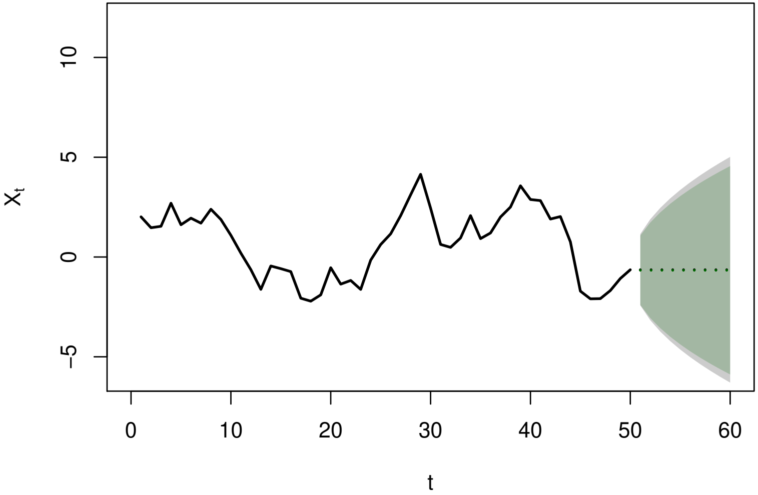
```
> modelfit <- robets(x)
> predict(modelfit,n.ahead=predtime)
```

- Using the ARMA fit

stats

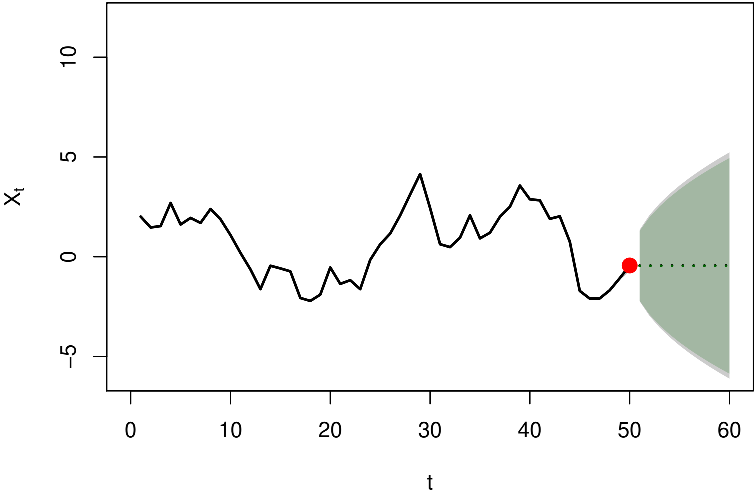
```
> armafit <- armarob(x,arorder=p,maorder=q,aic=TRUE)
> predict(armafit,n.ahead=predtime)
```

Sensitivity of forecasting by smoothing



Forecast of an AR(1) times series with parameter $\pi = 0.9$ with one outlier. Non robust forcast with ets (black) and robust forcast with robets (darkgreen).

Sensitivity of forecasting by smoothing



Forecast of an AR(1) times series with parameter $\pi = 0.9$ with one outlier. Non robust forecast with ets (black) and robust forcast with robets (darkgreen).

Exercises

1. Try to predict the number of covid cases in Austria for the next days.

References

Ruben Crevits and Christophe Croux. Forecasting using robust exponential smoothing with damped trend and seasonal components. *KBI_1741*, 2017.

H Dehling, R Fried, and M Wendler. A robust method for shift detection in time series. *Biometrika*, 107(3):647–660, 2020.

Alexander Dürre and Roland Fried. Robust change point tests by bounded transformations. *arXiv preprint arXiv:1905.06201*, 2019.

Alexander Dürre, Roland Fried, and Tobias Liboschik. Robust estimation of (partial) autocorrelation. *Wiley Interdisciplinary Reviews: Computational Statistics*, 7(3): 205–222, 2015.

Roland Fried. Robust filtering of time series with trends. *Journal of Nonparametric Statistics*, 16(3-4):313–328, 2004.

Edward J Hannan and Jorma Rissanen. Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, 69(1):81–94, 1982.

Yanyuan Ma and Marc G Genton. Highly robust estimation of the autocovariance function. *Journal of time series analysis*, 21(6):663–684, 2000.

Ricardo A Maronna, R Douglas Martin, and Victor J Yohai. *Robust statistics*. John Wiley & Sons, 2006.

J Mottonen, Visa Koivunen, and Hannu Oja. Robust autocovariance estimation based on sign and rank correlation coefficients. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics. SPW-HOS'99*, pages 187–190. IEEE, 1999.

References (cont.)

Bernhard Spangl. *On robust spectral density estimation*. PhD thesis, 2008.

Anita Monika Thieler, Michael Backes, Roland Fried, and Wolfgang Rhode. Periodicity detection in irregularly sampled light curves by robust regression and outlier detection. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 6(1):73–89, 2013.