COMPSTAT 2010

Robust forecasting of non-stationary time series*

Koen Mahieu

K.U.Leuven

* joint work with Christophe Croux, Roland Fried and Irène Gijbels

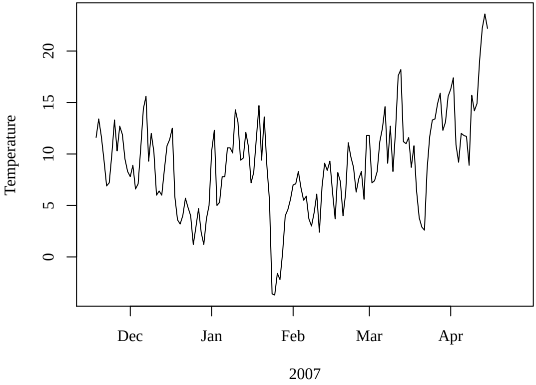
Goal

Forecasting of time series :

- non-stationarity
- non-parametric
- robust
- heteroscedasticity

Example

Temperature Data in Dresden, Germany



Idea

- non-stationarity $\longrightarrow$ local polynomial regression
- robust $\longrightarrow$ M-estimation
- heteroscedasticity $\longrightarrow$ MM-estimation

Model

$$Y_t = m(t) + \sigma(t)\epsilon_t,$$

Y_t	the observed time series,
$m(t)$	the signal,
$\sigma(t)$	the scale and
ϵ_t	the error term $\overset{\text{i.i.d}}{\sim} N(0,1)$.

Prediction

Based on the data $y_1, \ldots, y_T$, predict y_{t_0}, $t_0 > T$.

The signal is approximated locally by a polynomial of degree p:

$$m(t) = \sum_{j=0}^{p} \beta_j (t - t_0)^j,$$

where $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)'$ is to be estimated. Then

$$\hat{y}_{t_0} := \hat{m}(t_0) = \hat{\beta}_0.$$

Local polynomial regression

The OLS estimator of local polynomial regression for β minimizes

$$\sum_{t=1}^{T} \left(Y_t - \hat{\beta}' \mathbf{x}_{t,t_0} \right)^2 K \left(\frac{t - t_0}{h} \right),$$

with $\mathbf{x}_{t,t_0} = (1,(t-t_0),(t-t_0)^2,\ldots,(t-t_0)^p)'$, K an asymmetric kernel and h the bandwidth.

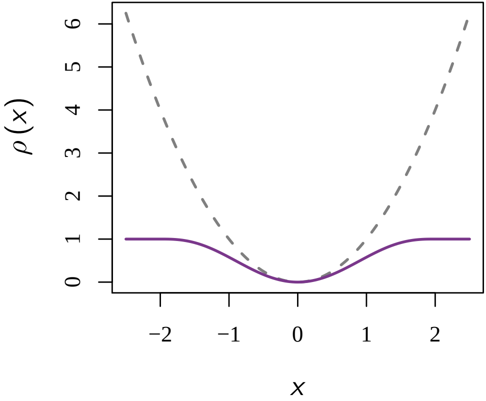
Local polynomial regression

The MM estimator of local polynomial regression for $\boldsymbol{\beta}$ minimizes

$$\sum_{t=1}^{T} \rho \left(\frac{Y_t - \hat{\boldsymbol{\beta}}' \mathbf{x}_{t,t_0}}{\hat{\sigma}(t_0)} \right) K \left(\frac{t - t_0}{h} \right),$$

where ρ is a loss function and $\hat{\sigma}(t_0)$ is the S-estimator of scale.

$$\left(Y_t - \hat{\boldsymbol{\beta}}' \mathbf{x}_{t,t_0}\right)^2 \rightarrow \rho\left(\frac{Y_t - \hat{\boldsymbol{\beta}}' \mathbf{x}_{t,t_0}}{\hat{\sigma}(t_0)}\right)$$



Algorithm

Iteratively weighted least squares:

- Weights = kernel weights × robustness weights
- β_{start} = Local least absolute deviation regression (LAD)
- σ_{start} = Locally weighted median absolute deviation from zero of the residuals of LAD-regression.

"Parameters to choose"

- degree of the polynomial ($p = 1$)
- kernel function ($K(x) = \exp(x) I_{\{x<0\}}$)
- breakdown point (50%)
- bandwidth h

Bandwidth selection

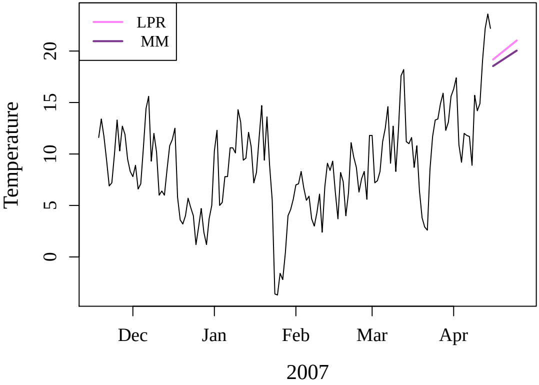
Select h such that the trimmed mean squared standardized forecast error is minimal:

$$h_{\text{opt}} = \operatorname*{argmin}_{h} \frac{1}{\lfloor 0.8T \rfloor} \sum_{i=1}^{\lfloor 0.8T \rfloor} \left(\frac{e_t}{\hat{\sigma}(t)} \right)^2_{(i)},$$

where $e_t/\hat{\sigma}(t)$ are the standardized forecast errors.

Example

Forecast Temperature Data



Example

Forecast accuracy

20% right trimmed means of the squared forecast errors of 50 one-step ahead forecasts.

	LPR	WRM	M	MM
TMSFE	7.37	8.31	7.03	6.88

Conclusion

Forecasting method for time series

- short-term
- non-parametric
- robust
- allow for non-stationarity and heteroscedasticity.

References

Croux, C., Gelper, S. and Mahieu, K. (2010).
Robust exponential smoothing of multivariate time series.
Computational Statistics & Data Analysis, **54**(12).

Grillenzoni, C. (2009).
Robust non-parametric smoothing of non-stationary time series.
Journal of Statistical Computation and Simulation, **79**(4).