Holt's exponential smoothing model for interval-valued time series

This work is part of a paper submitted to *International Journal of Forecasting*

André Luis Santiago Maia and Francisco de A. T. de Carvalho

Universidade Federal de Pernambuco
Centro de Informática
Av. Prof. Luiz Freire, s/n - Cidade Universitária
CEP: 50740-540 - Recife - PE - Brasil
`{alsm3,fatc}@cin.ufpe.br`

Contents

Symbolic Data Analysis

- Interval-valued data has been also considered in the field of *Symbolic Data Analysis* (SDA) (Bock and Diday (2000)).
- This field, related to multivariate analysis, pattern recognition and artificial intelligence, aims to extend classical exploratory data analysis and statistical methods to symbolic data.
- These new variables make it possible to take into account the variability and/or uncertainty present in the data.
- In the field of SDA, interval-valued data appear when the observed values of the variables are intervals of the set of real numbers $\mathbb{R}$.

Nowadays, different approches have been introduced to analyse interval-valued data.

- Patiño-Escarcina et al. (2004) propose a one layer perceptron for classification tasks, where inputs, weights and biases are represented by intervals.
- Roque et al. (2007) propose and analyse a new model of multilayer perceptron based on interval arithmetic that facilitates handling input and output interval data.
- Others authors have shown success with interval-valued data, but on interval analysis approach.

We manage interval-valued time series in the context of SDA, without use of operations and functions of interval arithmetic. This is a main feature that differs our paper from those cited above.

In the field of SDA for interval-valued data,

- Ichino et al. (1996) have introduced a symbolic classifier as a region oriented approach.
- Cazes et al. (1997) and Lauro and Palumbo (2000) introduced principal component analysis methods.
- Billard and Diday (2003) have introduced central tendency and dispersion measures.
- Chavent et al. (2006) and De Carvalho (2007) provides a number of clustering methods.
- Linear regression models have been also considered by Billard and Diday (2000) and Lima–Neto and De Carvalho (2008).
- Maia et al. (2008) propose approaches to interval-valued time series forecasting.

Contents

Interval-valued time series (ITS)

- When interval data is collected in an ordered sequence against time, we say that we have a *interval-valued time series*.
- The interval is described by a two-dimensional vector with elements in $\mathbb{R}$ represented by *upper bound*, X_t^U, and by *lower bound*, X_t^L.

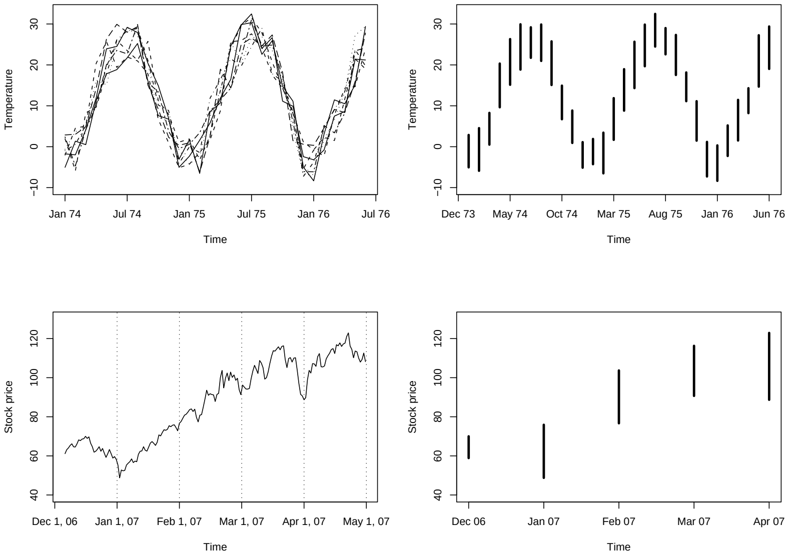
$$[X_1^L; X_1^U], [X_2^L; X_2^U], \ldots, [X_n^L; X_n^U]$$

- Specifically, an observed interval at time t is noted $\mathbf{I}_t$ and it is represented as

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix}$$

- Tools for interval-valued time series data analysis are also very much required.

Interval-valued time series (rigth) obtained from a set of classical time series (top left) and from time series of higher frequency (bottom left).



Motivation

Given an ITS, how to solve the problem of forecast in the context of SDA?

Contents

Standard Holt's method

In classical data, the standard Holt's method is given by

$$\begin{aligned} \widehat{L}_t &= \alpha y_t + (1-\alpha)(\widehat{L}_{t-1} + \widehat{T}_{t-1}), \\ \widehat{T}_t &= \beta(\widehat{L}_t - \widehat{L}_{t-1}) + (1-\beta)\widehat{T}_{t-1}. \end{aligned}$$

- $\widehat{L}_t$ is the smoothed level of the series, computed after y_t is observed
- $\widehat{T}_t$ is the smoothed trend at the end of period t
- $\widehat{y}_t = \widehat{L}_t + \widehat{T}_t$
- $0 < \alpha, \beta < 1$ are the smoothing parameters
- start values: $\widehat{L}_2 = y_2$ and $\widehat{T}_2 = y_2 - y_1$

Holt's method for ITS

The interval Holt's exponential smoothing method (Holt$^{\mathrm{I}}$) follows the representation

$$\begin{aligned}
\widehat{\mathbf{L}}_t^{\mathrm{I}} &= \mathcal{A}\mathbf{I}_t + (\mathrm{I} - \mathcal{A})(\widehat{\mathbf{L}}_{t-1}^{\mathrm{I}} + \widehat{\mathbf{T}}_{t-1}^{\mathrm{I}}), \\
\widehat{\mathbf{T}}_t^{\mathrm{I}} &= \mathcal{B}(\widehat{\mathbf{L}}_t^{\mathrm{I}} - \widehat{\mathbf{L}}_{t-1}^{\mathrm{I}}) + (\mathrm{I} - \mathcal{B})\widehat{\mathbf{T}}_{t-1}^{\mathrm{I}}.
\end{aligned}$$

$\mathcal{A}$ and $\mathcal{B}$ denote the (2×2) smoothing parameters matrices,

$$\mathcal{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}$$

and I is an (2×2) identity matrix.

Review

Standard Holt

$$\begin{aligned}\widehat{L}_t &= \alpha y_t + (1-\alpha)(\widehat{L}_{t-1} + \widehat{T}_{t-1}),\\ \widehat{T}_t &= \beta(\widehat{L}_t - \widehat{L}_{t-1}) + (1-\beta)\widehat{T}_{t-1}.\end{aligned}$$

$$\Downarrow$$

Interval Holt

$$\begin{aligned}\widehat{\mathbf{L}}^{\mathrm{I}}_t &= \mathcal{A}\mathbf{I}_t + (\mathrm{I}-\mathcal{A})(\widehat{\mathbf{L}}^{\mathrm{I}}_{t-1} + \widehat{\mathbf{T}}^{\mathrm{I}}_{t-1}),\\ \widehat{\mathbf{T}}^{\mathrm{I}}_t &= \mathcal{B}(\widehat{\mathbf{L}}^{\mathrm{I}}_t - \widehat{\mathbf{L}}^{\mathrm{I}}_{t-1}) + (\mathrm{I}-\mathcal{B})\widehat{\mathbf{T}}^{\mathrm{I}}_{t-1}.\end{aligned}$$

Expanding the expressions, the Holt^{I} method is given by:

$$\widehat{\mathbf{L}}_t^{\mathrm{I}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} + \begin{bmatrix} 1-\alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & 1-\alpha_{22} \end{bmatrix} \begin{bmatrix} \widehat{L}_{t-1}^U - \widehat{T}_{t-1}^U \\ \widehat{L}_{t-1}^L - \widehat{T}_{t-1}^L \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{11}X_t^U + (1-\alpha_{11})(\widehat{L}_{t-1}^U - \widehat{T}_{t-1}^U) + \alpha_{12}(X_t^L - \widehat{L}_{t-1}^L + \widehat{T}_{t-1}^L) \\ \alpha_{22}X_t^L + (1-\alpha_{22})(\widehat{L}_{t-1}^L - \widehat{T}_{t-1}^L) + \alpha_{21}(X_t^U - \widehat{L}_{t-1}^U + \widehat{T}_{t-1}^U) \end{bmatrix}$$

and

$$\widehat{\mathbf{T}}_t^{\mathrm{I}} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{L}_t^U - \widehat{L}_{t-1}^U \\ \widehat{L}_t^L - \widehat{L}_{t-1}^L \end{bmatrix} + \begin{bmatrix} 1-\beta_{11} & -\beta_{12} \\ -\beta_{21} & 1-\beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{T}_{t-1}^U \\ \widehat{T}_{t-1}^L \end{bmatrix}$$

$$= \begin{bmatrix} \beta_{11}(\widehat{L}_t^U - \widehat{L}_{t-1}^U) + (1-\beta_{11})\widehat{T}_{t-1}^U + \beta_{12}(\widehat{L}_t^L - \widehat{L}_{t-1}^L - \widehat{T}_{t-1}^L) \\ \beta_{22}(\widehat{L}_t^L - \widehat{L}_{t-1}^L) + (1-\beta_{22})\widehat{T}_{t-1}^L + \beta_{21}(\widehat{L}_t^U - \widehat{L}_{t-1}^U - \widehat{T}_{t-1}^U) \end{bmatrix}$$

The main advantage of the model presented here is the consideration that the trajectory of the upper boundary of the series can be affected by realizations of the lower boundary of this series and vice-versa.

$\mathcal{A}$ and $\mathcal{B}$ with elements constrained to the range $(0, 1)$ can be estimated by minimizing

$$
\begin{aligned}
\mathcal{R}(\mathcal{A},\mathcal{B}) &= \sum_{t=3}^{n} (\mathbf{I}_t - \widehat{\mathbf{I}}_t)^\top (\mathbf{I}_t - \widehat{\mathbf{I}}_t) \\
&= \sum_{t=3}^{n} \begin{bmatrix} X_t^U - \widehat{L}_{t-1}^U - \widehat{T}_{t-1}^U \\ X_t^L - \widehat{L}_{t-1}^L - \widehat{T}_{t-1}^L \end{bmatrix}^\top \begin{bmatrix} X_t^U - \widehat{L}_{t-1}^U - \widehat{T}_{t-1}^U \\ X_t^L - \widehat{L}_{t-1}^L - \widehat{T}_{t-1}^L \end{bmatrix} \\
&= \sum_{t=3}^{n} (X_t^U - \widehat{L}_{t-1}^U - \widehat{T}_{t-1}^U)^2 + \sum_{t=3}^{n} (X_t^L - \widehat{L}_{t-1}^L - \widehat{T}_{t-1}^L)^2.
\end{aligned}
$$

Start vectors: $\widehat{\mathbf{L}}_2^{\mathrm{I}} = \mathbf{I}_2$ and $\widehat{\mathbf{T}}_2^{\mathrm{I}} = \mathbf{I}_2 - \mathbf{I}_1$

Optimum smoothing parameters matrices for the Holt[I]

The estimation as a constrained non-linear programming problem:

$$\min_{\alpha_{ij},\beta_{ij}} \mathcal{R}(\mathcal{A},\mathcal{B}), \qquad \text{subject to} \qquad 0 \le \alpha_{ij}, \beta_{ij} \le 1$$

- The solution obtained by the *limited memory BFGS method for bound constrained optimization* (L-BFGS-B); Byrd et al. (1995)
- This method allows box constraints (each parameter can be given a lower and upper boundary)
- The L-BFGS-B algorithm is implemented in R software package; R Development Core Team (2008)

Contents

Application in stock market

Candlestick chart

http://finance.yahoo.com

- The applications in finance of the models are accomplished in interval-valued time series of stock market around days.[1]
- The time series used correspond to the stock prices, where the intervals are obtained for daily ranges:
  - the lowest traded price during the day (lower bound price, namely X^L)
  - the highest traded price during the day (upper bound price, namely X^U).

[1] Available in `http://finance.yahoo.com`

Tabela: Interval-valued time series processed.

Series	Period	Sample size
Itaú Holding	October 9, 2007 to March 17, 2008	109
Vale do Rio Doce	September 13, 2007 to March 13, 2008	126
Hollywood Media	July 9, 2007 to March 17, 2008	174
Petrobras	July 2, 2007 to March 13, 2008	177
Bradesco	April 3, 2007 to March 17, 2008	240
Brasil Telecom	January 10, 2007 to March 17, 2008	297
Google	August 10, 2006 to March 10, 2008	397
TAM	March 10, 2006 to March 17, 2008	507
Gol	December 13, 2005 to March 13, 2008	565
Apple	February 28, 2005 to March 14, 2008	767
Coca-cola	November 19, 2004 to March 17, 2008	834
Microsoft	February 18, 2003 to March 13, 2008	1277
Wal-Mart	February 25, 2000 to March 17, 2008	2024
GM	March 29, 1989 to March 17, 2008	4782
IBM	June 15, 1987 to March 17, 2008	5234

Performance measurements adapted for ITS

Interval U of Theil statistics

$$
\begin{aligned}
\mathrm{U}^{\mathrm{I}} &= \sqrt{\frac{\sum_{j=1}^{m}(\mathbf{I}_{j+1}-\widehat{\mathbf{I}}_{j+1})^{\top}(\mathbf{I}_{j+1}-\widehat{\mathbf{I}}_{j+1})}{\sum_{j=1}^{m}(\mathbf{I}_{j+1}-\mathbf{I}_{j})^{\top}(\mathbf{I}_{j+1}-\mathbf{I}_{j})}} \\
&= \sqrt{\frac{\sum_{j=1}^{m}(X_{j+1}^{U}-\widehat{X}_{j+1}^{U})^{2}+\sum_{j=1}^{m}(X_{j+1}^{L}-\widehat{X}_{j+1}^{L})^{2}}{\sum_{j=1}^{m}(X_{j+1}^{U}-X_{j}^{U})^{2}+\sum_{j=1}^{m}(X_{j+1}^{L}-X_{j}^{L})^{2}}}
\end{aligned}
$$

Results

	Interval U of Theil statistics (U^I)					
	Training set		5 steps ahead		10 steps ahead	
Series	Holt	$Holt^I$	Holt	$Holt^I$	Holt	$Holt^I$
Itaú Holding	1.003	0.962	2.193	0.871	2.426	0.913
Vale do Rio Doce	0.997	0.979	2.022	0.949	1.961	0.966
Hollywood Media	0.994	0.917	2.175	0.916	2.739	0.948
Petrobras	1.018	0.970	2.077	0.926	2.179	0.948
Bradesco	1.015	0.962	1.904	0.892	2.200	0.921
Brasil Telecom	0.995	0.945	1.107	0.936	1.111	0.939
Google	1.014	1.264	2.897	1.406	3.536	1.364
TAM	0.993	0.946	3.405	0.973	4.143	0.959
Gol	1.002	0.952	1.874	0.945	2.243	0.972
Apple	0.994	0.943	1.044	0.885	1.181	0.903
Coca-cola	1.002	0.944	1.059	0.962	1.449	0.902
Microsoft	0.987	0.953	1.426	0.931	1.461	0.932
Wal-Mart	1.014	0.946	2.373	0.936	2.530	0.937
GM	1.012	0.949	2.021	1.000	2.540	1.013
IBM	1.005	0.984	1.185	1.005	1.100	1.014
Mean	1.003	0.974	1.917	0.969	2.187	0.975
(St. dev.)	(0.010)	(0.082)	(0.682)	(0.127)	(0.875)	(0.113)

Contents

Conclusions

- We have proposed the Holt^I model based on Holt's exponential smoothing method
- The smoothing parameters are estimated by using techniques for non-linear optimization problems with box constraints (L-BFGS-B)
- The practicality of the methods is demonstrated by applications on real financial time series
- This method can be an alternative especially useful to stock prices modelling and forecasting
- The results suggest that the Holt^I model outperforms the standard Holt model
- Finally, our experiments suggest that this class of interval model can be successfully used for ITS

Main references

1. Bock, H.H. and Diday, E., *Analysis of Symbolic Data, Exploratory methods for extracting statistical information from complex data*. Springer, Heidelberg, 2000.

2. Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C., A limited memory algorithm for bound constrained optimization, *SIAM Journal Scientific Computing*, 16, 1190–1208, 1995.

3. Holt, C. C., Forecasting seasonals and trends by exponentially weighted averages. Reprinted with discussion in *International Journal of Forecasting*, 20, 5–13, 2004.

4. Maia, A. L. S., De Carvalho, F. A. T. and Ludermir, T. B., Forecasting models for interval-valued time series. *Neurocomputing*, 71, 3344–3352, 2008.

5. R Development Core Team , R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org`, 2008.