Variable Selection for long-term forecasting using temporal aggregation

Yves R. Sagaert, El-Houssaine Aghezzaf, Nikolaos Kourentzes, Bram Desmet

Department of Industrial Management, Ghent University

21/06/2016
Supply Chain Management constraints

Long term decisions require a good sales forecast

- Raw material
- Procurement negotiations
- Manufacturing and labor scheduling
- Capacity constraints
- Transportation
Traditional Sales Forecasting

Long term sales forecasting are formulated:

- Historical data patterns (level, trend, seasonality, ...)
- Promotions
- Judgemental adjustments:
  - Collaborative input from clients
  - Newspapers and industry magazines
  - Rumors in the corridors

Judgemental input is known to be biased and inconsistent (Fildes and Goodwin 2007, Trapero et al. 2013)

- Information of exogenous leading indicators
  - Capturing market sentiment in external big data (Russom et al. 2011)
The amount of newly registered cars (blue) is a leading indicator to the sudden drop (bold) in car tire sales (US) during the economic crisis of 2009-2010.
The curses of leading indicators

Curse of dimensionality

- Short fat data problem
- \( p > n \) : much more predictors than training sample

Curse of optimal leading effect

- Leading indicators exhibit leading information in advance
- \( pl \gg n \) : detecting optimal lead expands dimensionality

Curse of missing future information

- Indicators only exhibit information up to a certain point in time
- Clear need for unconditional forecasting
LASSO with limited sales history can improve on the company benchmark and on ETS

<table>
<thead>
<tr>
<th>Model</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive</td>
<td>17.205</td>
</tr>
<tr>
<td>Holt-Winters</td>
<td>18.590</td>
</tr>
<tr>
<td>Exponential Smoothing (ETS)</td>
<td>15.323</td>
</tr>
<tr>
<td>LASSO</td>
<td>13.781</td>
</tr>
</tbody>
</table>
Temporal aggregation

- Lower levels contain more noise and short term dynamics
- Cycles cannot be detected on lower frequency
- Capturing cycles is interesting for long term predictions
Temporal aggregation

- Indicator selected on lower levels contains more variance
- Indicator selected on high level is slower moving
Variable selection

- **Low level:**
  \[
  \hat{Y}_i = \beta_0 + \sum_{k=1}^{S} \beta_k D_k + \sum_{j=1}^{P} \beta_i x_{ij},
  \]
  seasonality selected on AIC

- **High level:**
  \[
  \hat{Y}_i = \beta_0 + \sum_{j=1}^{P} \beta_i x_{ij}
  \]

Forecast modeling on low level:

- Seasonality if selected
- Predictors selected on low/high level
Example results

The relative MAPE improvement between both variable selection methods

<table>
<thead>
<tr>
<th></th>
<th>1-6 months</th>
<th>6-12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low level</td>
<td>5 - 7 %</td>
<td></td>
</tr>
<tr>
<td>High level</td>
<td></td>
<td>1 - 6 %</td>
</tr>
</tbody>
</table>
Questions?

Thank you for your attention!

Yves Sagaert - yves.sagaert@ugent.be