Nowcasting with Dynamic Data Masking and Regularized Regression

Erik Alpkvist \(^1\) Jonas Hallgren \(^2\) Timo Koski \(^2\)
Johannes Siven \(^1\) Ard den Reijer \(^4\)

\(^1\)GSA Capital
London, UK

\(^2\)KTH, Royal Institute of Technology
Stockholm, Sweden

\(^4\)Sveriges Riksbank (Central Bank of Sweden)
Stockholm, Sweden

CFE 2015
Dec 12, 2015, University of London
In *nowcasting*, one wants to forecast a low frequency time series, $y$, using observations of several higher frequency time series, $Z$. A typical setting is:

\[ y \rightarrow \text{quarterly real GDP growth}. \]
\[ Z \rightarrow \text{a broad cross section of higher frequency economic indicators (e.g. weekly and monthly.)} \]

Task: Find $\mathbb{E}[y_N \mid Z_t]$. 
The image contains a graph with time series data from 1980 to 2015. The x-axis represents time, and the y-axis represents standardized observations. The graph includes the following data series:

- US GDP QoQ
- Nonfarm payrolls
- Chicago Purchasing Manager
- Durable Goods Orders

Each series is represented by a different line color or style.

The data points are distributed across the time period, showing variations in trends and patterns for each series.
DATA

- Release time stamps are essential.
- ALFRED is used in the study.
MODELS

- Models are on the form

\[ y_N = w^T F(Z_t) + \epsilon = w^T X + \epsilon. \]

We seek the projections \( F \).

- Two popular approaches: MIDAS-type and Dynamic Linear Factor models (DLM).

- The suggested approach is related to MIDAS.
MIDAS

- Mixed-data sampling (MIDAS)\(^1\) projects the time series onto each quarter

\[ F_{\text{MIDAS}}(Z_t) = X_{t}^{\text{MIDAS}} \]

by e.g. averaging the values. The prediction is

\[ \hat{y}_{N|t} = w^T X_{t}^{\text{MIDAS}}. \]

- Fine for historical regression.
- Forecasting is more difficult due to the “ragged edge” problem.

\(^1\)Ghysels et al. (2004)
MIDAS

\[ Z_{N-2} \rightarrow X_{N-2} \rightarrow y_{N-2} \]

\[ Z_{N-1} \rightarrow X_{N-1} \rightarrow y_{N-1} \]

\[ Z_t \rightarrow X_N \rightarrow y_N \]
DYNAMIC MASKING
Dynamic masking

- Simple idea: Construct regression features dynamically. Similar to MIDAS but the ragged edge problem disappears!
- Features are constructed for immediate use.
- The model is updated when data is updated.
- In both DLM and MIDAS the model is constant throughout the quarter and only data is updated.
### Algorithm

1. Find shape of the data available today.
3. Use MIDAS-type projection of masked data onto each quarter.
4. Regress on the masked data to get $w$.
5. Predict: $\hat{y}_{N|t} = w^T F_t^{\text{Mask}}$
ALGORITHM

1. Find shape of the data available today.
3. Use MIDAS-type projection of masked data onto each quarter.
4. Regress on the masked data to get $w$.
5. Predict: $\hat{y}_{N|t} = w^T F_t^{\text{Mask}} = w^T X_t^{\text{Mask}}$. 


**TIME t**

- $Z_{N-2} \rightarrow Z_{N-2}^M \rightarrow X_{N-2} \rightarrow y_{N-2}$
- $Z_{N-1} \rightarrow Z_{N-1}^M \rightarrow X_{N-1} \rightarrow y_{N-1}$
- $Z_t \rightarrow X_N \rightarrow y_N$
**TIME** $t + 1$

\[
\begin{align*}
Z_{N-2} & \rightarrow Z_{N-2}^{M_t+1} & \rightarrow X_{N-2} & \rightarrow y_{N-2} \\
Z_{N-1} & \rightarrow Z_{N-1}^{M_t+1} & \rightarrow X_{N-1} & \rightarrow y_{N-1} \\
Z_{t+1} & \rightarrow X_N & \rightarrow y_N
\end{align*}
\]
Why Masking?
Why Masking?

- Theoretical motivation
- Flexible and easy in implementation
- Good results
Time $t$

$p_\theta(\text{GDP}_N \mid \text{NFPAY}_{N-1}, \frac{\text{USD}}{\text{EUR}_t})$

In general: $\tilde{\theta} \neq \theta$!

---

Time $t + 1$

$p_{\tilde{\theta}}(\text{GDP}_N \mid \text{NFPAY}_N)$
**METHODS**

Given $X^{\text{Mask}}$ we predict $\hat{y} = w^T \Phi(X^{\text{Mask}})$.

- **Regression:** $\Phi$ is the identity and the estimates are
  - $w^{\text{OLS}} \leftarrow \arg\min_w \|y - w^T F^{\text{Mask}}\|_2$,
  - $w^{\text{Ridge}} \leftarrow \arg\min_w \|y - w^T F^{\text{Mask}}\|_2 + \lambda \|w\|_2$.

- **Neural Network**

$$\hat{y} = w^T \Phi^{\text{NN}}(X^{\text{Mask}}) = \sum_{k=1}^{M} w_k h(\theta_k^T X^{\text{Mask}}),$$

where $h$ is the activation function.

- **Kernel methods** predict using training data:

$$\Phi^{\text{Kernel}}(X^{\text{Mask}}) = k(X^{\text{Mask}}, X_{\text{Train}}).$$
Why Masking?

- Theoretical motivation
- Flexible and easy in implementation
- Good results
Any remark, question or suggestion is welcomed!

jonas@math.kth.se
http://people.kth.se/~jhallg/


Ragged Edge$^3$

1, 6 are published with lags. 2, 3, 5 are preliminary. 4 is given: interest rate.
Revision can take time:
In 2004 US revised the money supply (M2) series from January 1959 onwards$^2$!

$^2$Bouwman and Jacobs (2011)
$^3$Wallis (1986)