Forecasting with Weakly Identified Linear State-Space Models

Sébastien Blais

Discussion by
Marta Bańbura
European Central Bank

What the paper is about

Model: Gaussian linear state space model (LSSM):

$$\begin{aligned} \xi_{t+1} &= F\xi_t + v_t & \qquad v_t \sim N(0,Q) \\ y_t &= H'\xi_t + w_t & \qquad w_t \sim N(0,R) \end{aligned}$$

Discusses 2 problems:

- Normalisation
- Estimation and forecasting in case some of the parameters are weakly identified

Why do we need a normalisation?

... to overcome the problem of lack of identification.

- Likelihood: $l(\theta|y), \theta \in \Theta$
- It could happen that $\theta \neq \theta_1 \in \Theta$: $l(\theta|y) = l(\theta_1|y)$ for almost all possible $y \Rightarrow \theta$ and θ_1 *observationally equivalent*

Why can it be problematic?

- Consistent estimation
- Interpretation of the estimates
- Confidence bands and tests
- Computational problems

What is a normalisation?

- Normalisation: $\Theta^N \subset \Theta$

  e.g. $Y \sim N(0, \sigma^2)$: $l(\sigma|y) = l(-\sigma|y)$; Normalisation: $\sigma > 0$

- Model is identifiable (on Θ^N) if no two distinct $\theta, \theta_1 \in \Theta^N$ are observationally equivalent

  Often a weaker condition (*local indentification*) sufficient

Usually many normalisations possible, but ...

... it is not always innocuous how we choose to normalise

Some nice examples in Hamilton, Waggoner and Zha (2007).

Normalisation in this paper

- Some general remarks

  identification principle, discussion on the necessity and benefits of normalisation, discussion on the choice of priors in Bayesian framework

- Case of the LSSM:

  Parameters: $\theta = (H, F, R, Q) \in \Theta$

  θ not identified without further restrictions:

  M invertible matrix:

  $(\tilde{\theta}) = (\tilde{H}, \tilde{F}, \tilde{R}, \tilde{Q})$: $\tilde{H} = M'^{-1}H$, $\tilde{F} = MFM^{-1}$,...

  $l(\theta|Y) = l(\tilde{\theta}|Y)$

Normalisation in this paper, cntd

Again many possible normalisations.

Approach in this paper

- Decompose M into 4 “primitve” transformations (rescaling, rotation, permutation, reflection)
- Propose normalisations for each of these transformations separately (preserving the remaining invariance)

  Including for parametrisation in polar coordinates
- Propose how to implement the normalisations in Bayesian framework with Gibbs sampler

Weak identification

Situation close to lack of identification (and similar problems!)

- Concept not precisely defined in the literature
- Extensive literature in IV framework (weak instruments)

In the paper

- General definition not so clear (what is $\hat{\Theta}$?)
- In the empirical application:
  Factor (states) loadings are very small
- Example which shows that in case of weak identification adopting Bayesian approach results in better forecast than classical maximum likelihood

My impression

Interesting paper

- Potentially important question
- Not very much studied in the literature
- Novel approach

But ...

Comments: normalisation

- Studies only lack of identification cause by M transformation - there could be other issues, see below.
- Identification principle

  Theoretical grounds and/or practitioner's view?
- Some more discussion on the normalisations;

  Are all the 4 normalisations necessary for global identification?
- Normalisation of Harvey: is the problem relevant empirically?

  In the factor model literature usually dealt with by rank condition for H.
- What about other normalisations in the literature?
- Discussion on normalisation vs interpretation

Comments: weak identification

- Similar model studied by Stoffer and Wall (1991)

  Some discussion against their approach?
- Any other form of weak identification problem in LSSM?
- Is the problem relevant empirically?
- What is the role of the size of cross-section (N)?

See also later

Comments: normalisation vs weak identification

- Missing link
  - In Hamilton, Waggoner and Zha (2007) “reasonable” normalisation can alleviate the issues related to weak identification
  - Here the two not really connected (apart from the discussion)

    In the application normalisation is not imposed (??)
  - What is the connection?
  - Normalisation not so relevant for forecasting application

    Example of another application?
- Title

Relation to factor models

Factor model and LSSM are used interchangeably. However, while

- many forms of dynamic factor models can be cast in a state space form ...
- there is an important distinction (related to identification)

First identification issue in the factor model literature - distinguish common component (Λf_t) from the idiosyncratic (w_t):

- Covariance of w_t is diagonal
  relaxed for "large" models ($N \to \infty$), see e.g. Forni et al. (2000)
- $N \geq 2K + 1$
- Some version of full rank condition on Λ

see e.g. Heaton and Solo (2004).

Relation to factor models, cntd

To identify a particular "*M*-version" of the factors, some 0 restrictions imposed

(Often as in Harvey, 1989, see also Camba-Mendez et al. 2001, Jungbacker and Koopman, 2008)

However, Heaton and Solo (2004) show that these zero restrictions are not necessary if factors are sufficiently dynamically diversified.

(Some analogy with the result for ARMA processes as discussed in e.g. Harvey, 1989).

Coming back to the paper

- Important to acknowledge the distinction with respect to factor model assumptions
- Can we make connections?
  If so, in which situations are the normalisations proposed here useful?

Example of a weak identification problem in a factor model

$$
\begin{aligned}
y_t &= \Lambda f_t + w_t & \qquad w_t \sim N(0, R) \\
f_{t+1} &= F f_t + v_t & \qquad v_t \sim N(0, Q)
\end{aligned}
$$

Popular identification:

$$
\Lambda = \left[\frac{I_r}{\bar{\Lambda}} \right]
$$

Problem:
What if $y_{1,t}$ does not have a lot “in common” with the rest of the series (the loading is very small)

Solution:
Identify with a series on which we know that it strongly “comoves”

Conclusions

Interesting and novel paper.

Suggestions:

- Clearer connection of both considered issues
- Empirical relevance
- Appropriate connections to the factor model literature