# Discussion

*"Score-driven models for forecasting"*
by Siem Jan Koopman

**Domenico Giannone**
LUISS University of Rome, ECARES, EIEF and CEPR

8th ECB Forecasting Workshop

European Central Bank, June 2014

# Generalise Autoregressive Score (GAS) Basic Model Specification

$y_t$: data
$\hat{f}_t$: time varying parameter
$X_t$: exogenous variables
$\theta$: static parameters

$$y_t|\hat{f}_t \sim \tilde{p}(y_t|\hat{f}_t, \mathcal{F}_t; \theta)$$

$$\hat{f}_{t+1} = \omega + \sum_{i=1}^{p} A_i \hat{s}_{t-i+1} + \sum_{j=1}^{q} B_j \hat{f}_{t-j+1}$$

$$s_t = S_t \tilde{\nabla}_t$$

$$\tilde{\nabla}_t = \frac{\partial \log \tilde{p}(y_t|f_t; \Lambda)}{\partial f_t}$$

$$S_t = S(t, \hat{f}_t, \mathcal{F}_t; \theta)$$

Usually: $S_t = \alpha \mathcal{I}^{-1}$ where $\mathcal{I} = \mathrm{E}(\tilde{\nabla}_t \tilde{\nabla}'_t)$

# Generalise Autoregressive Score (GAS)

The origins:

- Creal, D., S. J. Koopman, and A. Lucas (2008). A general framework for observation driven time-varying parameter models. Discussion Paper 08-108/4, Tinbergen Institute.
- Harvey, A. C. and T. Chakravarty (2008). Beta-t-(E)GARCH. University of Cambridge, Faculty of Economics, Working paper CWPE 08340.

The presentation

- Introduction to GAS (Drew, Koopman, Lucas, 2013)
- Optimality of score (Blasques, Koopman and Lucas, 2014)
- Forecasting evidence (Koopman, Lucas and Scharth, 2014)

# My Quest for the Origins of GAS

- Siem Jan Koopman:
  Time series analysis by **state space** methods
- Andrew C. Harvey:
  Forecasting, structural time series models and the **Kalman filter**

The discussion

- From State Space Models to Score Driven Models
- Spot the difference
  - Dynamic Factor Models (DFM)
    *Traditional vs Score Driven*

# Linear State space model

$$y_t | f_t \sim \mathcal{N}(\Lambda f_t, R)$$
$$f_{t+1} | f_t \sim \mathcal{N}(A f_t, Q)$$

Defining $\hat{f}_{t+1} := \mathrm{E}(f_{t+1} | y^t)$, we have the Kalman recursion:

$$\hat{f}_{t+1} = A\hat{f}_t + AG_t v_t$$

where

- $v_t = (y_t - \Lambda \hat{f}_t)$: surprises
- $G_t = P_t \Lambda' (\Lambda P_t \Lambda' + R)^{-1}$: gain

Where the variance of the estimated states is computed from the recursion:

$$\mathrm{E}[(\hat{f}_{t+1} - f_{t+1})(\hat{f}_{t+1} - f_{t+1})'] := P_{t+1} = AP_t A' + Q - AP_t K_t \Lambda'$$

# Linear State Space model

$$y_t | f_t \sim \mathcal{N}(\Lambda f_t, R)$$

$$f_{t+1} | f_t \sim \mathcal{N}(A f_t, Q)$$

$$\hat{f}_{t+1} = A\hat{f}_t + A G_t v_t$$

It can be shown that:

$$G_t v_t = (\underbrace{\Lambda' R^{-1} \Lambda}_{\mathcal{I}} + P_t^{-1})^{-1} \underbrace{\Lambda' R^{-1} v_t}_{\nabla_t}$$

$$\nabla_t = \frac{\partial \log p(y_t | f_t; \Lambda)}{\partial f_t}, \quad \mathcal{I} = \mathrm{E}(\nabla_t \nabla_t')$$

# Kalman and GAS

**Linear State Space model** (Parameter-driven)

$$y_t | f_t \sim \mathcal{N}(\Lambda f_t, R)$$

$$f_{t+1} | f_t \sim \mathcal{N}(A f_t, Q)$$

$$\hat{f}_{t+1} = A\hat{f}_t + A(\mathcal{I}_t + P_t^{-1})^{-1} \nabla_t$$

$$\nabla_t = \frac{\partial \log p(y_t | f_t; \Lambda)}{\partial f_t}, \quad \mathcal{I} = \mathrm{E}(\nabla_t \nabla_t')$$

**Generalised Autoregressive Scores (GAS)** (Observation-driven)

$$y_t | \hat{f}_t \sim \tilde{p}(\hat{f}_t; \theta)$$

$$\hat{f}_{t+1} = A\hat{f}_t + \alpha(\tilde{\mathcal{I}}_t + 0)^{-1} \tilde{\nabla}_t$$

$$\tilde{\nabla}_t = \frac{\partial \log \tilde{p}(y_t | f_t; \Lambda)}{\partial f_t}, \quad \mathcal{I} = \mathrm{E}(\tilde{\nabla}_t \tilde{\nabla}_t')$$

# Kalman and GAS

**Non-Linear State space model** (Parameter-driven)

$$y_t|f_t \sim p_{y|f}(y_t|f_t;\theta)$$

$$f_{t+1}|\hat{f}_t \sim p_{f_+|f}(f_{t+1}|f_t;\theta)$$

**Generalised Autoregressive Scores (GAS)**(Observation-driven)

$$y_t|\hat{f}_t \sim \tilde{p}(y_t|\hat{f}_t;\theta)$$
$$\hat{f}_{t+1} = A\hat{f}_t + \alpha(\ \tilde{\mathcal{I}}_t\ )^{-1}\tilde{\nabla}_t$$

$$\tilde{\nabla}_t = \frac{\partial \log \tilde{p}(y_t|f_t;\Lambda)}{\partial f_t}, \quad \mathcal{I} = \mathrm{E}(\tilde{\nabla}_t\tilde{\nabla}_t')$$

# Dynamic Factor Model (DFM)

**Traditional DFM**(Engle and Watson, 1983 ...)

- $y_t = \Lambda f_t + e_t \sim \mathcal{N}(0, R)$
- $f_{t+1} = A f_t + u_t \sim \mathcal{N}(0, Q)$

$$\hat{f}_{t+1} = A\hat{f}_t + AG_t(y_t - \Lambda \hat{f}_t)$$

$$\mathrm{E}[(\hat{f}_{t+1} - f_{t+1})(\hat{f}_{t+1} - f_{t+1})'] := P_{t+1}$$

**Obs.-driven DFM** (Creal, Schwaab, Koopman and Lucas, 2014)

- $y_t = \Lambda \hat{f}_t + e_t \sim \mathcal{N}(0, R)$
- $\hat{f}_{t+1} = A\hat{f}_t + \alpha(\Lambda' R^{-1} \Lambda)^{-1} \Lambda' R^{-1}(y_t - \Lambda \hat{f}_t)$

$\Longrightarrow \hat{f}_{t+1} = A\hat{f}_t + \alpha(\hat{\hat{f}}_t - \hat{f}_t)$ where $\hat{\hat{f}}_t = (\Lambda' R^{-1} \Lambda)^{-1} \Lambda' R^{-1} y_t$

$$\text{if } \alpha = A \text{ then } \hat{f}_{f+1} = A\hat{\hat{f}}_t$$

# Conclusions

- Interesting, Flexible and Useful Approximation of Linear State Space Models
- Interesting, Flexible and Useful Approximation of Non-Linear State Space Models
- Very intriguing and Interesting Research Agenda
- Open Questions (at least for me)
  - Hard to interpret in some instance
  - When is the approximation reasonable? How does it compare with other approximations (e.g. forgetting factor, Koop and Korobolis, 2013)
  - Does it work in forecasting? (See Opschoor et al, 2014)
  - Are computational gains going to remain relevant?
- Strongly suggest to read the papers