# On the consistency of the MLE for observation-driven models

Paolo Gorgi

Joint work with:
F. Blasques
S. J. Koopman

Università degli Studi di Padova

# Outline

- In observation-driven models **invertibility is needed** to
  1) Ensure the consistency of the MLE.
  2) Uncover the true path of the time varying parameter (even if $\theta_0$ is known).

- **Problem: existing conditions for invertibility are often useless in practice**. In particular, to ensure invertibility to hold we need to impose severe restrictions that are unreasonable in empirical applications.

- **Solution:** we derive the consistency of the MLE considering feasible invertibility conditions that can cover situations of practical interest.

# Motivation: the model



- Consider the **Beta-t-GARCH model** with leverage effects of Creal et al. (2013) and Harvey (2013)

$$\begin{aligned} y_t &= \sqrt{f_t}\varepsilon_t, \quad \varepsilon_t \sim t_v(0,1), \\ f_{t+1} &= \omega + \beta f_t + (\alpha + \gamma d_t)\frac{(v+1)y_t^2}{(v-2) + f_t^{-1} y_t^2}, \end{aligned}$$

where $d_t = 1$ if $y_t \leq 0$ and $d_t = 0$ otherwise.

- To ensure the consistency of the MLE, the parameter region $\Theta$ where the likelihood is maximized has to satisfy

$$E \log \left| \beta + (\alpha + \gamma d_t) \frac{(v+1)y_t^4}{\left((v-2)\bar{\omega} + y_t^2\right)^2} \right| < 0, \quad \forall\, \theta \in \Theta,$$

and $\theta_0 \in \Theta$.

# Motivation: the parameter region



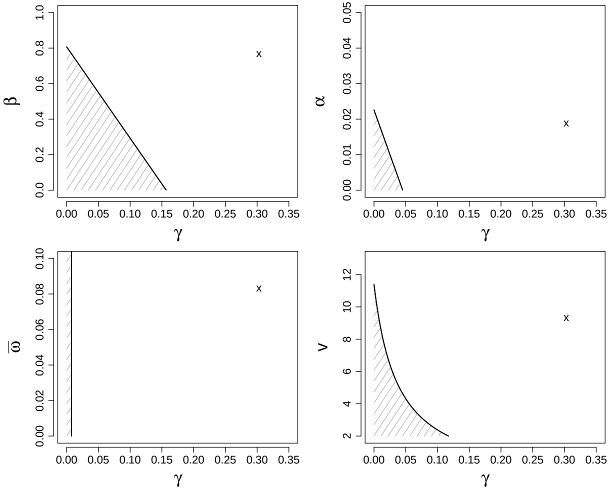


Figure: Parameter region where we can ensure that the invertibility condition hold. The cross denotes the parameter estimate using monthly log-differences of the S&P 500 stock index.

# Observation-driven models



- We observe data $\{y_t\}_{t=1}^{n}$, and we consider the following model

$$
\begin{aligned}
y_t|f_t &\sim p(y_t|f_t,\theta),\\
f_{t+1} &= \phi(f_t, y_t, \theta),\ t \in \mathbb{Z},
\end{aligned}
$$

  where $p(\cdot|f_t;\theta)$ is a density function, $\theta \in \Theta$ a parameter vector and $\phi$ is a continuous function.

- Under the assumption of correct specification, the data generating process (DGP) satisfies the model equations at $\theta = \theta_0$ and $f_t^o$ denotes the true time varying parameter.

- We are interested in ML estimation of the static parameter $\theta$ and, in particular, the consistency of the MLE.

# The likelihood function



- Using the observed data, the **filtered parameter** is obtained as

$$\hat{f}_{t+1}(\theta) = \phi(\hat{f}_t(\theta), y_t, \theta), \quad t \in \mathbb{N},$$

  for an **initial value** $\hat{f}_1(\theta) \in \mathcal{F}_\theta \subseteq \mathbb{R}$.

- The MLE is then obtained maximizing the likelihood

$$\hat{L}_n(\theta) = n^{-1} \sum_{t=1}^{n} \log p(y_t | \hat{f}_t(\theta), \theta),$$

  over the parameter set $\Theta$.

- The stability (**invertibility**) of $\{\hat{f}_t(\theta)\}_{t \in \mathbb{N}}$ for the $\theta \in \Theta$ plays a key role to ensure the consistency of the MLE.

# Invertibility



- The filtered parameter $\{\hat{f}_t(\theta)\}_{t\in\mathbb{N}}$ at $\theta$ is invertible if

$$\left|\hat{f}_t(\theta) - \tilde{f}_t(\theta)\right| \xrightarrow{a.s.} 0, \qquad \text{as} \;\; t \to \infty.$$

  for any $\hat{f}_1(\theta) \in \mathcal{F}$, where $\{\tilde{f}_t(\theta)\}_{t\in\mathbb{Z}}$ is a stochastic sequence.

- Invertibility guarantees that the path of the true time varying parameter $f_t^o$ can be recovered asymptotically, i.e. $|\hat{f}_t(\theta_0) - f_t^o| \xrightarrow{a.s.} 0$.

**Invertibility is not merely a technical condition**, see Sorokin (2011) and Wintenberger (2013).

# Why is invertibility important?

**EGARCH(1,1):** $y_t = \exp(f_t/2)\varepsilon_t, \quad f_{t+1} = \omega + \beta f_t + \alpha|\varepsilon_t|.$

- $|\beta_0| < 1$ **ensures stationarity** of the EGARCH(1,1) process.
- $|\beta_0| < 1$ **does not ensure invertibility** of the filter $\hat{f}_t(\theta_0)$.

Plot of $10^{-5}\sum_{t=1}^{10^5} |f_t^o - \hat{f}_t(\theta_0)|$ for different initializations $\hat{f}_1(\theta_0)$.

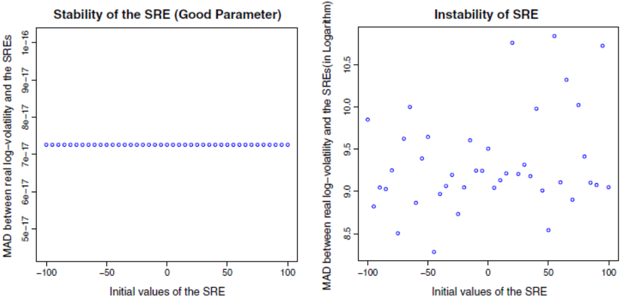


Figure: Non-invertibility example of EGARCH(1,1) from Wintenberger (2013).

# How can we ensure invertibility?



- As in Straumann and Mikosh (2006), **sufficient conditions** for invertibility can be obtained on the basis of Theorem 3.1 of Bougerol (1993).

- Bougerol's theorem provides general conditions for stability of stochastic processes.

- We obtain that $\{\hat{f}_t(\theta)\}_{t\in\mathbb{N}}$ is invertible if

$$\boldsymbol{E \log \Lambda_t(\theta) < 0},$$

where

$$\Lambda_t(\theta) = \sup_f \left| \frac{\partial \phi(f, y_t, \theta)}{\partial f} \right|.$$

# Invertibility: GARCH and EGARCH



- **GARCH(1,1) model**

  Filtered parameter:
  $\hat{f}_{t+1}(\theta) = \omega + \beta \hat{f}_t(\theta) + \alpha y_t^2.$

  Invertibility condition:
  $E \log \sup_f |\partial(\beta f + \alpha y_t^2)/\partial f| = \mathbf{\log(\beta)} < \mathbf{0}.$

- **EGARCH(1,1) model**

  Filtered parameter:
  $\hat{f}_{t+1}(\theta) = \omega + \beta \hat{f}_t(\theta) + \alpha |y_t| \exp\left(-\hat{f}_t(\theta)/2\right).$

  Invertibility condition:
  $E \log \sup_f |\partial\left(\beta f + \alpha |y_t| \exp(-f/2)\right)/\partial f| =$
  $\boldsymbol{E \log \max\left\{\beta, 2^{-1}\alpha |y_t| \exp(-2^{-1}\omega/(1-\beta)) - \beta\right\} < 0}.$

# Invertibility in practice



- Often, in practice, $E \log \Lambda_t(\theta) < 0$ **cannot be checked directly** as $\Lambda_t(\theta)$ depends on the unknown data generating process.
- This leads to either a very small region or a degenerate region $\Theta$ where the likelihood should maximized

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} \hat{L}_n(\theta),$$

- In practical applications, invertibility conditions are ignored and therefore the **consistency of the MLE is not guaranteed** and it may not be possible to uncover the true path $f_t^o$.

# MLE on an empirical region

UNIVERSITÀ DEGLI STUDI DI PADOVA

- To handle this issue we define the MLE on a parameter region that satisfies an **empirical** version of the **invertibility condition** $E \log \Lambda_t(\theta) < 0$, namely

$$\tilde{\theta}_n = \arg\max_{\theta \in \hat{\Theta}_n} \hat{L}_n(\theta),$$

where

$$\hat{\Theta}_n = \left\{ \theta \in \bar{\Theta} : \frac{1}{n} \sum_{t=1}^{n} \log \Lambda_t(\theta) < 0 \right\}.$$

- Wintenberger (2013) first proposed the estimation of the parameter region for the QMLE of the EGARCH(1,1) model.

# Consistency of the MLE



We consider the following conditions:

(C.1) The data generating process is stationary with $E \log \Lambda_t(\theta_0) < 0$.

(C.2) The model is identifiable.

(C.3) The $\log \Lambda_t(\theta)$ is a.s. continuous and it has a finite first moment.

(C.4) The log-likelihood function is Lipschitz continuous with respect to $\hat{f}_t(\theta)$.

(C.5) The first moment of the likelihood function is uniformly bounded.

**Theorem**

*Let conditions (C.1)-(C.5) hold, then the MLE $\tilde{\theta}_n$ is strongly consistent, i.e.*

$$\tilde{\theta}_n \xrightarrow{a.s.} \theta_0, \qquad n \to \infty.$$

*Furthermore, $|\hat{f}_n(\tilde{\theta}_n) - f_n^o| \xrightarrow{a.s.} 0$ as $n$ goes to infinity.*

# Example 1: the model



- The **Beta-t-GARCH model** with leverage effects of Creal et al. (2013) and Harvey (2013) is

$$
\begin{aligned}
y_t &= \sqrt{f_t}\varepsilon_t, \quad \varepsilon_t \sim t_v(0,1), \\
f_{t+1} &= \omega + \beta f_t + (\alpha + \gamma d_t)\frac{(v+1)y_t^2}{(v-2) + f_t^{-1} y_t^2},
\end{aligned}
$$

  where $d_t = 1$ if $y_t \leq 0$ and $d_t = 0$ otherwise.

- The invertibility condition $E \log \Lambda_t(\theta) < 0$ is given by

$$
E \log \left| \beta + (\alpha + \gamma d_t)\frac{(v+1)y_t^4}{\left((v-2)\bar{\omega} + y_t^2\right)^2} \right| < 0, \quad \forall\, \theta \in \Theta.
$$

# Example 1: the parameter region

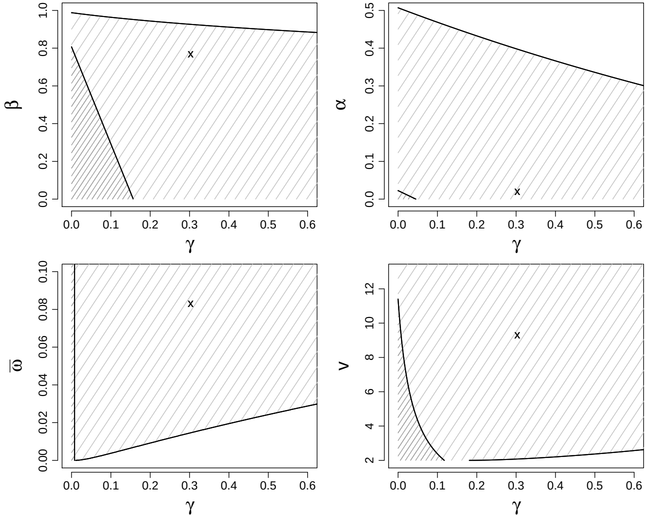


Figure: Invertibility regions obtained considering the monthly log-differences of the S&P 500 stock index.

# Example 2: the model



- The **dynamic autoregressive model** Blasques et al. (2014) and Delle Monache and Petrella (2014) is

$$y_t = f_t y_{t-1} + \sigma \varepsilon_t, \quad \varepsilon_t \sim t_v,$$
$$f_{t+1} = \omega + \beta f_t + \alpha \frac{(y_t - f_t y_{t-1}) y_{t-1}}{1 + v^{-1}\sigma^{-2}(y_t - f_t y_{t-1})^2},$$

- The invertibility condition $E \log \Lambda_t(\theta) < 0$ is given by

$$E \log \max \left\{ \left| \beta - \alpha y_t^2 \right|, \left| \beta + \frac{\alpha}{8} y_t^2 \right| \right\} < 0, \quad \forall\, \theta \in \Theta.$$

- Sufficient conditions leads to a degenerate region with $\alpha = 0$.

# Example 2: the parameter region



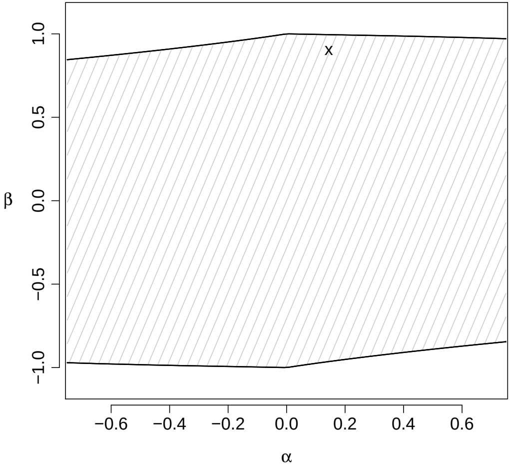


Figure: Invertibility region obtained considering the monthly log-differences of the US unemployment claims.

# Concluding remarks



- Even if $\theta_0$ does not satisfy $E \log \Lambda_0(\theta_0) < 0$, the MLE $\tilde{\theta}_n$ and the filtered parameter $\hat{f}_n(\tilde{\theta}_n)$ asymptotically does not depend on the initialization $\hat{f}_1(\theta)$.

- In the case of model misspecification, the MLE $\tilde{\theta}_n$ is consistent w.r.t. a pseudo true parameter $\theta^*$. This pseudo true parameter has the interpretation of being the minimizer of the following marginal KL divergence

$$KL(\theta) = E \log p^o(y_t|y^{t-1}) - E \log p(y_t|f_t(\theta), \theta),$$

where $p^o(y_t|y^{t-1})$ denotes the unknown true conditional distribution of $y_t$.

# Main references



Bougerol, P. (1993). Kalman filtering with random coefficients and contractions. *Control Optim.* **31**, 942–959.

Straumann, D. & Mikosch, T. (2006). Quasi maximum likelihood estimation in conditionally heteroschedastic time series: a stochastic recurrence equation approach. *Ann. Statist.* **34**, 2449–2495.

Wintenberger, O. (2013). Continuous Invertibility and Stable QML Estimation of the EGARCH(1,1) Model. *Scandinavian Journal of Statistics* **40**, 846–867.