Advanced Time Series Assignment

Wiktor Grocholewski

June 2025

SVAR(1) with 2 variables proof of LAN

Statistical Model:

$$(\Omega, \mathscr{F}, \{\mathbb{P} : \theta \in \Theta = \mathbb{R}^6 || \Gamma| \neq 0 \wedge \rho(\Gamma^{-1}B) < 1\}$$

for:

$$\theta = [\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}]'$$

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \tag{1}$$

Which can be written as:

$$\Gamma Y_t = BY_{t-1} + \varepsilon_t$$

With $\mathbb{E}[\varepsilon_t] = 0$, and $Var(\varepsilon_t) = \Sigma_\varepsilon < \infty$

Likelihood ratio:

First I write down the likelihood:

$$\mathbb{P}_\theta(Y_t | Y_1, ..., Y_{t-1}) = f(\Gamma Y_t - BY_{t-1})$$

Where f is the density of a vector ε_t, that is smooth in θ.
Now I check the likelihood ratio of the local experiment of $\theta_T = \theta_0 + \frac{h}{\alpha_T}$ vs $\theta_T = \theta_0$:

$$log \frac{d\mathbb{P}_{\theta_T}}{d\mathbb{P}_{\theta_0}} = \sum_{t=0}^{T} log \frac{f(\Gamma_T Y_t - B_T Y_{t-1})}{f(\Gamma_0 Y_t - B_0 Y_{t-1})} = \sum_{t=0}^{T} log \frac{f(\varepsilon_t + (\Gamma_T - \Gamma_0)Y_t - (B_T - B_0)Y_{t-1})}{f(\varepsilon_t)} \approx$$

$$\approx \sum_{t=0}^{T} log(1 + (\theta_T - \theta_0)' \frac{1}{f(\varepsilon_t)} \frac{\partial \varepsilon_t'}{\partial \theta} \nabla_\varepsilon f(\varepsilon_t) \approx \frac{h}{\alpha_T} \sum_{t=0}^{T} \frac{1}{f(\varepsilon_t)} \frac{\partial \varepsilon_t'}{\partial \theta} \nabla_\varepsilon f(\varepsilon_t) - \frac{1}{2\alpha_T^2} \sum_{t=0}^{T} h' \frac{1}{f(\varepsilon_t)^2} \frac{\partial \varepsilon_t'}{\partial \theta} \nabla_\varepsilon f(\varepsilon_t) \nabla_\varepsilon f(\varepsilon_t)' (\frac{\partial \varepsilon_t'}{\partial \theta})' h =$$

$$= h' \Delta_T - \frac{1}{2} h' I_T h$$

for:

$$\Delta_T = \frac{1}{\alpha_T} \sum_{t=0}^{T} \frac{1}{f(\varepsilon_t)} \frac{\partial \varepsilon_t'}{\partial \theta} \nabla_\varepsilon f(\varepsilon_t)$$

$$I_T = \frac{1}{\alpha_T^2}\sum_{t=0}^{T}\frac{1}{f(\varepsilon_t)^2}\frac{\partial\varepsilon_t'}{\partial\theta}\nabla_\varepsilon f(\varepsilon_t)\nabla_\varepsilon f(\varepsilon_t)'(\frac{\partial\varepsilon_t'}{\partial\theta})'$$

Here we need to note that:

$$\mathbb{E}[\frac{1}{f(\varepsilon_t)}\frac{\partial\varepsilon_t'}{\partial\theta}\nabla_\varepsilon f(\varepsilon_t)]$$

Which is just a transformation of ε_t, so if enough moments exist and since we imposed stationarity, should be eligible for applying WLLN and CLT after substracting the Expected value, under $\alpha_T = \sqrt{T}$. Also the Fisher Information Matrix should be finite.

What would happen if we couldn't observe y_{2t}:

The obvious consequence is that since we have less information, the Fisher Information will be lower, and as a result the variance will be bigger.

Integrating out y_{2t}:

The goal right now is to integrate out the unobserved varialbe y_{2t}, so the likelihood function would be:

$$\int \mathbb{P}_\theta(Y_t|Y_1,...,Y_{t-1})dy_{2*} = \int f(\Gamma Y_t - BY_{t-1})dy_{2*} = g(y_{1t}|y_{11},..,y_{1t-1})$$

In this section I will keep this form, but I will try to describe the y_{1t} process in the next. To make my life easier I will assume that the errors follow a normal distribution, i.e. $\varepsilon_t \sim N(0,\Sigma_\varepsilon)$, with $\Sigma_\varepsilon < \infty$. Under this more restrictive assumption the resulting distribution after integrating out y_{2*} should also be gaussian. With that the likelihood ratio of the local experiment should be:

$$log\int\frac{d\mathbb{P}_{\theta_T}}{d\mathbb{P}_{\theta_0}}(Y_t|Y_1,...,Y_{t-1})dy_{2*} = \sum_{t=0}^{T} log\frac{\int f(\Gamma_T Y_t - B_T Y_{t-1})dy_{2*}}{\int f(\Gamma_0 Y_t - B_0 Y_{t-1})dy_{2*}} =$$

$$= \sum_{t=0}^{T} log\frac{g_T(y_{1t}|y_{11},..,y_{1t-1})}{g_0(y_{1t}|y_{11},..,y_{1t-1})} = h'\,\mathbb{E}[\Delta_T|y_{1*}] - \frac{1}{2}h'\,\mathbb{E}[I_T|y_{1*}]h$$

Since I end up with the likelihood ratio of the two normal distributions this is basically the same as the Gaussian Experiment and so this distritubion satisfies LAN as long as we assume stationarity and existence of enough moments which I did already for the model with both variables observed.

Regressing just y_{1t}:

To see what would happen if I only decided to regress y_{1t}, I tried to substitute in y_{2t} into y_{1t}, and I ended up with an ARMA(2,1) process. Here I show the steps how I got there:
(1) I rewrite the SVAR(1) into VAR(1):

$$\Gamma Y_t = BY_{t-1} + \varepsilon_t => Y_t = \Gamma^{-1}BY_{t-1} + \Gamma^{-1}\varepsilon_t = AY_{t-1} + \Gamma^{-1}\varepsilon_t$$

(2) With that I can write it as a system of two equations:

$$\begin{cases} y_{1t} = a_{11}y_{1t-1} + a_{12}y_{2t-1} + g_{11}\varepsilon_{1t} + g_{12}\varepsilon_{2t} \\ y_{2t} = a_{21}y_{1t-1} + a_{22}y_{2t-1} + g_{21}\varepsilon_{1t} + g_{22}\varepsilon_{2t} \end{cases} \tag{2}$$

(3) I will substitute y_{2t} into y_{1t}:

$$y_{1t} = a_{11}y_{1t-1} + a_{12}[a_{21}y_{1t-2} + a_{22}y_{2t-2} + g_{21}\varepsilon_{1t-1} + g_{22}\varepsilon_{2t-1}] + g_{11}\varepsilon_{1t} + g_{12}\varepsilon_{2t}$$

(4) Now I need to substitute in y_{2t-2} which I can take from the equation for y_{1t}:

$$y_{2t-1} = \frac{1}{a_{12}}(y_{1t} - a_{11}y_{1t-1} - g_{11}\varepsilon_{1t} - g_{12}\varepsilon_{2t})$$

$$y_{1t} = a_{11}y_{1t-1}+a_{12}[a_{21}y_{1t-2}+\frac{a_{22}}{a_{12}}(y_{1t}-a_{11}y_{1t-1}-g_{11}\varepsilon_{1t}-g_{12}\varepsilon_{2t})+g_{21}\varepsilon_{1t-1}+g_{22}\varepsilon_{2t-1}]+g_{11}\varepsilon_{1t}+g_{12}\varepsilon_{2t}$$

(5) Which I can rewrite as:

$$y_{1t} = \frac{1}{1-a_{22}}(a_{11}(1-a_{22})y_{1t-1} + a_{12}a_{21}y_{1t-2} + (1-a_{22})[g_{11}\varepsilon_{1t} + g_{12}\varepsilon_{2t}] + a_{12}[g_{21}\varepsilon_{1t} + g_{22}\varepsilon_{2t}]$$

From that we know that when y_{2t} is unobserved, the process for y_{2t} is ARMA(2,1). That being said this is weaker as Fisher Information will be lower, and therefore the true parameters will be harder to identify.