Multivariate LSTAR (in Geodesy)

Tomáš Bacigál

Department of Mathematics and Descriptive Geometry

Faculty of Civil Engineering

Bratislava

Outline

- Multivariate threshold autoregressive model.
- Multivariate logistic smooth transition AR.
- Testing for (non)linearity.
- Model building.
- Application.

Multivariate Threshold AR model

Consider a k-dimensional time series

$$\boldsymbol{y}_t = (y_{1t}, \ldots, y_{kt})'$$

and v-dimensional exogenous variables

$$\boldsymbol{x}_t = (x_{1t}, \ldots, x_{vt})' .$$

Then $\boldsymbol{y}_t$ follows a multivariate threshold model with threshold variable z_t and delay d if it satisfies

$$\boldsymbol{y}_t = \begin{cases} \boldsymbol{\phi}_0^{(1)} + \sum_{i=1}^{p} \boldsymbol{\phi}_i^{(1)} \boldsymbol{y}_{t-i} + \sum_{i=1}^{q} \boldsymbol{\beta}_i^{(1)} \boldsymbol{x}_{t-i} + \varepsilon_t^{(1)} & \text{if } z_{t-d} \leq c \\ \boldsymbol{\phi}_0^{(2)} + \sum_{i=1}^{p} \boldsymbol{\phi}_i^{(2)} \boldsymbol{y}_{t-i} + \sum_{i=1}^{q} \boldsymbol{\beta}_i^{(2)} \boldsymbol{x}_{t-i} + \varepsilon_t^{(2)} & \text{if } z_{t-d} > c \end{cases}$$

or more general

$$\boldsymbol{y}_t = \boldsymbol{\phi}_0^{(j)} + \sum_{i=1}^{p} \boldsymbol{\phi}_i^{(j)} \boldsymbol{y}_{t-i} + \sum_{i=1}^{q} \boldsymbol{\beta}_i^{(j)} \boldsymbol{x}_{t-i} + \varepsilon_t^{(j)} \quad \text{if} \;\; c_{j-1} < z_{t-d} \leq c_j \, , \tag{1}$$

where c is a threshold value and $j = 1, \dots, s$ is index of regime. The threshold variable z_t is assumed to be stationary.

The model (1) has s regimes and is piecewise linear in the threshold space z_{t-d}, but it is nonlinear in time when $s > 1$.

The transition between different regimes is sharp.

Multivariate Smooth Transition AR

A more graded transition between the different regimes can be obtained by replacing the brittle function $1(c < z_{t-d})$ by a continuous transition function $G(z_{t-d}, \gamma, c)$ that changes smoothly from 0 to 1 as y_{t-d} increases. The resulting model

$$\begin{aligned}\boldsymbol{y}_t = &\left(\boldsymbol{\phi}_0^{(1)} + \sum_{i=1}^{p} \boldsymbol{\phi}_i^{(1)} \boldsymbol{y}_{t-i} + \sum_{i=1}^{q} \boldsymbol{\beta}_i^{(1)} \boldsymbol{x}_{t-i}\right)(1 - G(z_{t-d}, \gamma, c)) \\ &+ \left(\boldsymbol{\phi}_0^{(2)} + \sum_{i=1}^{p} \boldsymbol{\phi}_i^{(2)} \boldsymbol{y}_{t-i} + \sum_{i=1}^{q} \boldsymbol{\beta}_i^{(2)} \boldsymbol{x}_{t-i}\right) G(z_{t-d}, \gamma, c) + \varepsilon_t \qquad (2)\end{aligned}$$

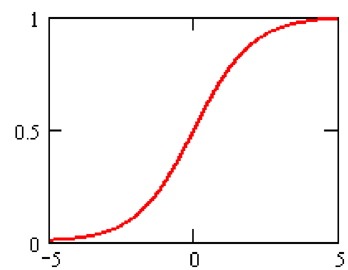
is called Smooth Transition AR (STAR).

... Logistic Smooth Transition AR

A popular choice for the transition function is the logistic function

$$G(z_{t-d}, \gamma, c) = \frac{1}{1 + e^{-\gamma(z_{t-d} - c)}}$$

and the resulting model is called Logistic STAR (LSTAR) model.



The parameter c can be interpreted as the *threshold* between the two regimes, and parameter γ determines the smoothness of the transition.

Testing for nonlinearity

Given observations $\{\boldsymbol{y}_t, \boldsymbol{x}_t, z_t\}$, for $t = 1, \ldots, n$, the goal is to detect the threshold nonlinearity of $\boldsymbol{y}_t$.

If the null hypothesis ($\boldsymbol{y}_t$ is linear) holds, then the OLS estimates of

$$\boldsymbol{y}_t' = \boldsymbol{X}_t' \boldsymbol{\Phi} + \boldsymbol{\varepsilon}_t', \quad t = h+1, \ldots, n$$

are useful.

In the regression, $\boldsymbol{X}_t = (1, \boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}', \boldsymbol{x}_{t-1}', \ldots, \boldsymbol{x}_{t-q}')'$ is a $(pk + qv + 1)$-dimensional regressor, $h = max(p, q, d)$, and $\boldsymbol{\Phi}$ denotes parameter matrix.

If the regression is rearranged according to increasing order of the threshold variable z_{t-d}, then the threshold model is effectively transformed into a changepoint problem, i.e.,

$$\boldsymbol{y}'_{t(i)+d} = \boldsymbol{X}'_{t(i)+d}\boldsymbol{\Phi} + \boldsymbol{\varepsilon}'_{t(i)+d}\,, \quad t = 1, \ldots, n-h\,, \tag{3}$$

where $t(i)$ is the time index of i-th smallest z.

To detect model change in (3), predictive residuals and recursive least squares method can be used.

Let $\hat{\boldsymbol{\Phi}}_m$ be the LS-estimate of $\boldsymbol{\Phi}$ with $i = 1, \ldots, m$. Let

$$\hat{\boldsymbol{e}}_{t(m+1)+d} = \boldsymbol{y}_{t(m+1)+d} - \hat{\boldsymbol{\Phi}}'_m \hat{\boldsymbol{X}}_{t(m+1)+d}$$

and

$$\hat{\boldsymbol{\eta}}_{t(m+1)+d} = \hat{\boldsymbol{e}}_{t(m+1)+d} / [1 + \hat{\boldsymbol{X}}'_{t(m+1)+d} \boldsymbol{V}_m \hat{\boldsymbol{X}}_{t(m+1)+d}]^{1/2},$$

where $\boldsymbol{V}_m = [\sum_{i=1}^{m} \hat{\boldsymbol{X}}'_{t(i)+d} \hat{\boldsymbol{X}}'_{t(i)+d}]^{-1}$, be the predictive and the standardized predictive residual of regresion (3).

Next, consider the regression

$$\hat{\boldsymbol{\eta}}'_{t(l)+d} = \boldsymbol{X}'_{t(l)+d}\boldsymbol{\Psi} + \boldsymbol{w}'_{t(l)+d}\,, \quad l = m_0 + 1, \ldots, n - h\,,$$

where m_0 denotes starting point of the recursive LS-estimation.
The problem of interest is then to test the hypothesis $H_0 :\ \boldsymbol{\Psi} = 0$ versus the alternative $H_1 :\ \boldsymbol{\Psi} \neq 0$.
There is a test statistic

$$C(d) = [n - h - m_0 - (kp + vq + 1)]\,(\ln[det(S_0)] - \ln[det(S_1)])\ ,$$

where the delay d signifies, that the test depends on the threshold variable z_{t-d}, and

$$S_0 = \frac{1}{n-h-m_0} \sum_{l=m_0+1}^{n-h} \hat{\boldsymbol{\eta}}_{t(l)+d} \hat{\boldsymbol{\eta}}'_{t(l)+d} \,,$$

and

$$S_1 = \frac{1}{n-h-m_0} \sum_{l=m_0+1}^{n-h} \hat{\boldsymbol{w}}_{t(l)+d} \hat{\boldsymbol{w}}'_{t(l)+d}$$

Under the null hypothesis that $\boldsymbol{y}_t$ is linear and some regularity conditions, $C(d)$ is asymptotically a χ^2 random variable with $k(pk+qv+1)$ degrees of freedom.

Model building

- Identification of the appropriate threshold value c, lag order d and orders p, q by minimization of information criterion.
- Estimation of the parameters $\hat{\boldsymbol{\theta}} = (\boldsymbol{\phi}^{(1)}, \boldsymbol{\phi}^{(2)}, \gamma, c)'$ in the STAR model by straightforward application of nonlinear least squares

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} \sum_{t=1}^{n} \left[\boldsymbol{y}_t - F(\boldsymbol{X}_t; \boldsymbol{\theta})\right]^2 ,$$

where

$$F(\boldsymbol{X}_t; \boldsymbol{\theta}) = \boldsymbol{\phi}^{(1)} \boldsymbol{X}_t [1 - G(\boldsymbol{y}_{t-d}; \gamma, c)] + \boldsymbol{\phi}^{(2)} \boldsymbol{X}_t G(\boldsymbol{y}_{t-d}; \gamma, c) .$$

Application

thank you