Ekonometria Finansowa

Nieliniowe modele szeregów czasowych

mgr Paweł Jamer[1]

Doktorant, Katedra Ekonometrii i Statystyki SGGW
Ekspert ds. Modelowania Danych, Polskie Technologie
Konsultant Zewnętrzny, Polkomtel

7 października 2016

[1] pawel.jamer@gmail.com

Problem

Problem

Poszukiwany jest model opisujący szereg czasowy r_t równaniem postaci

$$r_t = f\left(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right),$$

gdzie

- f — dowolna funkcja nieliniowa,
- ϵ_t — $\mathrm{WN}\left(0, 1\right)$

Uproszczenie problemu

Uprośćmy przedstawiony problem rozwijając funkcję f w szereg Taylora:

$$r_t = g\left(\epsilon_{t-1}, \epsilon_{t-2}, \ldots\right) + \epsilon_t h\left(\epsilon_{t-1}, \epsilon_{t-2}, \ldots\right).$$

Wówczas funkcje g oraz h możemy interpretować następująco:

- $g\left(\epsilon_{t-1}, \epsilon_{t-2}, \ldots\right) = \mathbb{E}\left(r_t \mid \mathcal{R}_{t-1}\right),$
- $h^2\left(\epsilon_{t-1}, \epsilon_{t-2}, \ldots\right) = \mathrm{Var}\left(r_t \mid \mathcal{R}_{t-1}\right).$

Definicja

Model nieliniowy

Model opisujący proces r_t nazwiemy nieliniowym, jeżeli warunkowa wartość oczekiwana $\mathbb{E}\left(r_t \mid \mathcal{R}_{t-1}\right)$ lub warunkowa wariancja $\operatorname{Var}\left(r_t \mid \mathcal{R}_{t-1}\right)$ jest w nim opisana za pomocą funkcji nieliniowej.

Model nieliniowej autoregresji

Modelem nieliniowej autoregresji rzędu p szeregu czasowego r_t nazwiemy model opisany równaniem

$$y_t = f\left(y_{t-1}, y_{t-2}, \ldots, y_{t-p}\right) + \epsilon_t.$$

Oznaczenie. Model nieliniowej autoregresji rzędu p oznacza się symbolem $\mathrm{NLAR}\,(p)$.

Model wykładniczej autoregresji

Modelem wykładniczej autoregresji rzędu p szeregu czasowego r_t nazwiemy model opisany równaniem

$$y_t = \sum_{j=1}^{p} \alpha_j y_{t-j} + \epsilon_t,$$

gdzie

- $\alpha_j = \theta_j + \pi_j e^{-\alpha y_{t-j}^2}.$

Oznaczenie. Model wykładniczej autoregresji rzędu p oznacza się symbolem $\mathrm{EAR}(p)$.

Model nieliniowej średniej ruchomej

Modelem nieliniowej średniej ruchomej rzędu q szeregu czasowego r_t nazwiemy model opisany równaniem

$$y_t = f\left(\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-q}\right) + \epsilon_t.$$

Oznaczenie. Model nieliniowej średniej ruchomej rzędu q oznacza się symbolem $\mathrm{NLMA}\left(q\right)$.

Model Rocke'a

Modelem Rocke'a szeregu czasowego r_t nazwiemy model opisany równaniem

$$y_t = \epsilon_t + \sum_{j=1}^{\infty} b_j \psi\left(\epsilon_{t-j}\right),$$

gdzie

- $\psi\left(x\right) \leqslant 0$ dla $x > 0$,
- $\psi\left(-x\right) = \psi\left(x\right)$,
- $\psi\left(0\right) = 1$.

Model dwuliniowy

Modelem dwuliniowym szeregu czasowego r_t nazwiemy model opisany równaniem

$$y_t = \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=0}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{P} \sum_{l=1}^{Q} c_{k,l} y_{t-k} \epsilon_{t-l},$$

gdzie

- $\theta_0 = 1$.

Uwaga. Model dwuliniowy uwzględnia skupianie się danyh.

Uwaga. $r_t \sim \text{ARCH} \Rightarrow r_t^2$ - proces dwuliniowy.

Wśród modeli dwuliniowych wyróżnia się:

- modele dwuliniowe naddiagonalne ($c_{k,l} = 0$ dla $k > l$).
- modele dwuliniowe poddiagonalne ($c_{k,l} = 0$ dla $k < l$).
- modele dwuliniowe diagonalne ($c_{k,l} = 0$ dla $k \neq l$).

Model TAR

Model progowy autoregresyjny

Modelem progowym autoregresyjnym (TAR) szeregu czasowego r_t nazwiemy model opisany równaniem

$$r_t + \theta_0 + \sum_{i=1}^{p} \theta_i r_{t-i} + I\left(z_{t-d} \leqslant r\right)\left(\psi_0 + \sum_{j=1}^{q} \psi_j r_{t-j}\right) = \epsilon_t.$$

Model SETAR. Model TAR w którym $z_{t-d} = r_{t-d}$ nazwiemy modelem SETAR.

Model wygładzonego przejścia

Modelem wygładzonego przejścia szeregu czasowego r_t nazwiemy model opisany równaniem

$$r_t = \boldsymbol{x}'_t \boldsymbol{\pi}_1 + \boldsymbol{x}'_t \boldsymbol{\pi}_2 F(z_t) + \epsilon_t,$$

gdzie

- $\boldsymbol{x}_t = \begin{bmatrix} 1 & y_{t-1} & \cdots & y_{t-p} & x_{1,t} & \cdots & x_{q,t} \end{bmatrix}$ — wektor zmiennych,
- $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$ — wektory parametrów,
- ϵ_t — WN,
- $F(z_t)$ — funkcja transformacji, parzysta lub nieparzysta funkcja ciągła.

Funkcja logistyczna:

$$F(z_t) = F(t) = \frac{1}{1 + e^{-\gamma\left(t^k + \sum_{i=1}^{k} \alpha_i t^{k-i}\right)}}.$$

Funkcja wykładnicza:

$$F(z_t) = F(t) = 1 + e^{-\gamma(t-\alpha)^2}.$$

Model przełącznikowy

Model przełącznikowy

Modelem przełącznikowym szeregu czasowego r_t nazwiemy model opisany równaniem

$$r_t = \alpha_{z_t} + \sum_{i=1}^{p} \beta_{z_t,i} r_{t-i} + \epsilon_t,$$

gdzie

- z_t — łańcuch Markowa o zbiorze stanów

$$\mathcal{S} = \{1, 2, \ldots, k\}, k \geqslant 2.$$

Odpowiednio dobrana funkcja nieliniowa powinna:

- zachowywać zgodność ze znanymi zależnościami ekonomicznymi,
- zachowywać zgodność ze znanymi faktami empirycznymi,
- posiadać dziedzinę zmienności zgodną z faktyczną zmiennością modelowanej wielkości,
- charakteryzować się elastycznością umożliwiająca aproksymację form zbliżonych do analizowanej formy.
- umożliwiać łatwe wyznaczenie jej parametrów.

Współczynnik maksymalnej korelacji:

$$m_\rho = \max_{f,g} \operatorname{Cor}\left(g\left(y\right), f\left(x\right)\right).$$

Maksymalna średnia korelacja:

$$m_m = \max_{f} \operatorname{Cor}\left(y, f\left(x\right)\right).$$

Maksymalny współczynnik regresji

$$m_r = \max_{f} R^2,$$

gdzie

- R^2 — współczynnik determinacji w regresji $y = f\left(x\right) + \epsilon_t$.

Wybór liczby opóźnień

Uwaga. Wyboru liczby opóźnień w modelach NLAR, NLMA, dwuliniowych oraz progowych można dokonać wykorzystując kryteria informacyjne. Najczęściej stosowanymi kryteriami są kryteria AIC oraz BIC.

1. Stosujemy MNK w celu wyznaczenia parametrów modelu
$$y_t = \boldsymbol{\beta}' \mathbf{z}_t + u_t,$$
gdzie $\mathbf{z}_t = \begin{bmatrix} 1 & y_{t-1} & \cdots & y_{t-p} & x_1 & \cdots & x_k \end{bmatrix}$.
2. Obliczamy
$$\mathrm{SSR}_0 = \sum_{t=1}^{T} \hat{u}_t^2.$$
3. Stosujemy MNK w celu wyznaczenia parametrów modelu
$$\hat{u}_t = \boldsymbol{\delta}' \mathbf{z}_t + \sum_{j=2}^{h} \psi_j \hat{y}_t^j + v_t.$$
4. Obliczamy
$$\mathrm{SSR} = \sum_{t=1}^{T} \hat{v}_t^2.$$

Testujemy hipotezę

$$\begin{cases} H_0: & (\forall j)\, \psi_j = 0, \\ H_1: & (\exists j)\, \psi_j \neq 0. \end{cases}$$

Statystyka testowa jest postaci:

$$F = \frac{(\mathrm{SSR}_0 - \mathrm{SSR})/(h-1)}{(\mathrm{SSR})/(T-m-h)} \sim \mathbb{F}^{[h-1,\, T-m-h]},$$

gdzie $m = p + k$.

Pytania?

Dziękuję za uwagę!