

Ekonometria Finansowa

Jednowymiarowe modele szeregów czasowych

mgr Paweł Jamer¹

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

25 października 2016

¹pawel.jamer@gmail.com

Proces AR

Zdefiniujmy operator

$$\varphi(B) = I - \varphi_1 B - \dots - \varphi_p B^p,$$

gdzie $p \in \mathbb{Z}_+$.

Proces AR

Słabo stacjonarny szereg czasowy X_t nazwiemy procesem AR (autoregresyjnym) rzędu p , jeżeli spełnia on równanie

$$\varphi(B) X_t = \epsilon_t,$$

gdzie $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Oznaczenie. Proces AR rzędu p oznacza się symbolem $\text{AR}(p)$.

Proces MA

Zdefiniujmy operator

$$\theta(B) = I + \theta_1 B + \dots + \theta_q B^q,$$

gdzie $q \in \mathbb{Z}_+$.

Proces MA

Słabo stacjonarny szereg czasowy X_t nazwiemy procesem MA (średniej ruchomej) rzędu q , jeżeli spełnia on równanie

$$X_t = \theta(B) \epsilon_t,$$

gdzie $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Oznaczenie. Proces MA rzędu q oznacza się symbolem $\text{MA}(q)$.

Proces ARMA

Definicja

Proces ARMA

Słabo stacjonarny szereg czasowy X_t nazwiemy procesem ARMA (p, q) , jeżeli spełnia on równanie

$$\varphi(B) X_t = \theta(B) \epsilon_t,$$

gdzie $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Proces ARMA

Estymowanie funkcji autokorelacji (ACF)

Estymator funkcji autokorelacji

Jako estymator funkcji autokorelacji procesu X_t możemy przyjąć

$$\hat{\rho}_h = \frac{\sum_{t=1}^{T-h} (X_t - \bar{X}) (X_{t+h} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

Proces ARMA

Estymowanie funkcji autokorelacji częściowej (PACF)

Estymator funkcji autokorelacji częściowej

Jako estymator funkcji autokorelacji częściowej procesu X_t możemy przyjąć

$$\alpha_0 = 1,$$

$$\alpha_1 = \phi_{1,1} = \hat{\rho}_1,$$

$$\alpha_h = \phi_{h,h} = \frac{\hat{\rho}_h - \sum_{j=1}^{h-1} \phi_{h-1,j} \hat{\rho}_{h-j}}{1 - \sum_{j=1}^{h-1} \phi_{h-1,j} \hat{\rho}_j}, h = 2, 3, \dots,$$

gdzie $\phi_{i,j}$ to rozwiązania układu Yule'a-Walkera.

Intuicja. Współczynnik autokorelacji częściowej mierzy korelację między zmiennymi X_t oraz X_{t-h} po eliminacji wpływu zmiennych pośrednich.

Proces ARMA

Identyfikacja na podstawie ACF oraz PACF

- Jeżeli proces jest typu $AR(p)$, to funkcja autokorelacji powoli maleje, natomiast funkcja autokorelacji częściowej staje się statystycznie równa zero od wartości $p + 1$.
- Jeżeli proces jest typu $MA(q)$, to funkcja autokorelacji staje się statystycznie równa zero od wartości $q + 1$, natomiast funkcja autokorelacji częściowej powoli maleje.
- Jeżeli proces jest typu $ARMA(p, q)$, to funkcja autokorelacji oraz funkcja autokorelacji częściowej łagodnie zanikają.

Proces ARMA

Identyfikacja na podstawie kryteriów informacyjnych

Idea

Należy oszacować modele ARMA (p, q) dla wszystkich możliwych kombinacji parametrów $p = 1, 2, \dots, P$ oraz $q = 1, 2, \dots, Q$, a następnie obliczyć dla nich wartości wybranego kryterium informacyjnego. Minimalna wartość kryterium wyznacza optymalną parę parametrów p i q .

Proces ARMA

Kryteria informacyjne

Niech $L_{p,q}$ oznacza wiarygodność modelu ARMA (p, q) .

Kryterium informacyjne Akaike (AIC)

$$\text{AIC}_{p,q} = -2 \ln L_{p,q} + 2(p + q + 1)$$

Bayesowskie kryterium informacyjne (BIC)

$$\text{BIC}_{p,q} = -2 \ln L_{p,q} + (p + q + 1) \ln T$$

Proces ARMA

Podział metod estymacji

Metody estymacji:

- Metody estymacji parametrów modelu AR (p) :
 - równania Yule'a-Walkera,
 - metoda najmniejszych kwadratów,
 - metoda największej wiarygodności.
- Metody estymacji parametrów modelu ARMA (p, q) :
 - metoda największej wiarygodności.

Proces ARMA

Idea metod estymacji

Równania Yule'a-Walkera

Równania Yule'a-Walkera to liniowy układ równań pozwalający na wyrażenie parametrów autoregresji za pomocą współczynników autokorelacji. Oszacowania parametrów modelu $AR(p)$ otrzymujemy zastępując we wspomnianym układzie współczynniki autokorelacji ich estymatorami, a następnie rozwiązując układ.

Metoda najmniejszych kwadratów

Wykorzystanie klasycznej MNK do stacjonarnego szeregu czasowego typu $AR(p)$. Uzysane w ten sposób estymatory będą zgodne z estymatorami metody największej wiarygodności.

Metoda największej wiarygodności

Metoda polegająca na maksymalizacji funkcji największej wiarygodności procesu $ARMA(p, q)$.

Proces ARMA

Prognoza

Interesuje nas prognoza na h okresów w przód:

$$X_{t+h} = \sum_{k=1}^p \rho_k X_{t+h-k} + \sum_{k=0}^q \theta_k \epsilon_{t-k},$$

gdzie $\theta_0 = 1$.

Możemy wyrazić ją jako warunkową wartość oczekiwaną:

$$\begin{aligned}\hat{X}_t(h) &= \mathbb{E}(X_{t+h} \mid X_t, X_{t-1}, \dots), \\ \mathbb{E}(X_{t+j} \mid X_t, X_{t-1}, \dots) &= \begin{cases} X_{t+j}, & j \leq 0, \\ \hat{X}_t(j), & j > 0, \end{cases} \\ \mathbb{E}(\epsilon_{t+j} \mid \epsilon_t, \epsilon_{t-1}, \dots) &= \begin{cases} \epsilon_{t+j}, & j \leq 0, \\ 0, & j > 0. \end{cases}\end{aligned}$$

Proces ARIMA

Proces ARIMA

Szereg czasowy X_t nazwiemy procesem ARIMA (p, d, q) , jeżeli szereg czasowy $\Delta^d X_t$ jest procesem ARMA (p, q) .

Z powyższej definicji wynika, że proces ARIMA (p, d, q) jest opisywany przez następujące równanie:

$$\varphi(B) (\Delta^d X_t) = \theta(B) \epsilon_t.$$

Multiplikatywny proces ARMA

Intuicja

Niech proces X_t charakteryzuje równanie

$$\varphi_X(B) X_t = \theta_X(B) \epsilon_t,$$

natomiast proces ϵ_t równanie

$$\varphi_\epsilon(B) \epsilon_t = \theta_\epsilon(B) \eta_t,$$

gdzie $\eta_t \sim \text{WN}(0, \sigma^2)$. Składając powyższe równania uzyskujemy

$$\varphi_\epsilon(B) \varphi_X(B) X_t = \theta_X(B) \theta_\epsilon(B) \eta_t.$$

Uwaga. Powyższe złożenie dwóch procesów ARMA pozostaje nadal procesem ARMA.

Multiplikatywny proces ARMA

Definicja

Multiplikatywny proces ARMA

Szereg czasowy X_t nazwiemy multiplikatywnym procesem ARMA $(p, q) \times (P, Q)_s$, jeżeli spełnia on równanie

$$\Phi(B^s)X_t = \Theta(B^s)\epsilon_t,$$

gdzie

- Φ — operator analogiczny do φ rzędu P ,
- Θ — operator analogiczny do θ rzędu Q ,
- ϵ_t — proces ARMA (p, q) .

Multiplikatywny proces ARMA

Właściwości

Reprezentacja. Proces $\text{ARMA}(p, q) \times (P, Q)_s$ jest specyficznym przypadkiem procesu $\text{ARMA}(p + sP, q + sQ)$.

Estymacja. Estymacja parametrów multiplikatywnych procesów ARMA odbywa się przez osobną estymację parametrów każdego z procesów ARMA wchodzących w jego skład.

Sezonowy proces ARIMA

Sezonowy proces ARIMA

Szereg czasowy X_t nazwiemy sezonowym procesem ARIMA $(p, d, q) \times (P, D, Q)_s$, jeżeli spełnia on równanie

$$\Phi(B^s)\phi(B)\Delta_s^D\Delta^dX_t = \Theta(B^s)\theta(B)\epsilon_t,$$

gdzie

- ϵ_t — biały szum,
- $\Delta_s^D\Delta^dX_t$ — proces stacjonarny.

Stylizowane fakty

Stylizowane fakty dla stóp zwrotu:

- Rozkłady stóp zwrotu mają grubsze ogony niż rozkład normalny.
- Wartości stóp zwrotu są nieskorelowane, ale ich kwadraty są skorelowane.
- Duże zmiany wartości stóp zwrotu następują często po wcześniejszych dużych zmianach.

Model ARCH

Model ARCH

Powiemy, że szereg czasowy r_t jest procesem ARCH(p), jeżeli

$$\begin{aligned}r_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^p \alpha_j r_{t-j}^2,\end{aligned}$$

gdzie

- ϵ_t iid WN(0, 1),
- $\alpha_0 > 0$,
- $\alpha_j \geq 0$ dla $j = 1, 2, \dots, p$.

Uwaga. Model ARCH jest często wykorzystywany do opisu reszt w modelach ARMA.

Stacjonarność

Stacjonarność

Stacjonarne szeregi czasowe typu ARCH (p) spełniają nierówność

$$\sum_{j=1}^p \alpha_j < 1.$$

Warunkowa wariancja

Niech \mathcal{R}_τ będzie zbiorem wszystkich informacji o szeregu czasowym r_t do chwili czasu τ . Wówczas:

$$\text{Var}(r_t | \mathcal{R}_{t-1}) = \mathbb{E}(r_t^2 | \mathcal{R}_{t-1}) - \mathbb{E}^2(r_t | \mathcal{R}_{t-1}),$$

$$\text{Var}(r_t | \mathcal{R}_{t-1}) = \mathbb{E}(\sigma_t^2 \epsilon_t^2 | \mathcal{R}_{t-1}) - \mathbb{E}^2(\sigma_t \epsilon_t | \mathcal{R}_{t-1}),$$

$$\text{Var}(r_t | \mathcal{R}_{t-1}) = \sigma_t^2 \mathbb{E}(\epsilon_t^2 | \mathcal{R}_{t-1}) - \sigma_t^2 \mathbb{E}^2(\epsilon_t | \mathcal{R}_{t-1}),$$

$$\text{Var}(r_t | \mathcal{R}_{t-1}) = \sigma_t^2 \mathbb{E}(\epsilon_t^2) - \sigma_t^2 \mathbb{E}^2(\epsilon_t),$$

$$\text{Var}(r_t | \mathcal{R}_{t-1}) = \sigma_t^2.$$

Wniosek. σ_t^2 jest warunkową wariancją procesu r_t pod warunkiem przeszłości \mathcal{R}_{t-1} .

Efekt ARCH

Test Engle'a

Szacujemy parametry modelu:

$$r_t^2 = \alpha_0 + \sum_{j=1}^p \alpha_j r_{t-j}^2 + \eta_t, \eta_t \text{ iid WN}(0, s).$$

Testujemy hipotezę postaci:

$$H_0 : \sum_{j=1}^p \alpha_j^2 = 0.$$

Statystyka testowa przyjmuje postać:

$$LM = TR^2 \rightarrow \chi_p^2,$$

gdzie R^2 to współczynnik determinacji modelu wyjściowego.

Efekt ARCH

Test McLeod-Li

Badamy szereg czasowy kwadratów wartości szeregu czasowego wyjściowego r_t .

Testujemy hipotezę postaci

$$H_0 : \sum_{j=1}^h \rho_j^2 = 0.$$

Statystyka testowa przyjmuje postać:

$$T_{LB} = n(n+2) \sum_{j=1}^h \frac{\hat{\rho}_j^2}{n-j} \rightarrow \chi_h^2,$$

gdzie

$$\hat{\rho}_j = \frac{\sum_{t=h+1}^T r_t^2 r_{t-h}^2}{\sum_{t=1}^T r_t^4}.$$

Model GARCH

Model GARCH

Powiemy, że szereg czasowy r_t jest procesem GARCH (p, q) , jeżeli

$$\begin{aligned}r_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^q \alpha_j r_{t-j}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,\end{aligned}$$

gdzie

- ϵ_t iid WN $(0, 1)$,
- $\alpha_0 > 0$,
- $\alpha_j \geq 0$ dla $j = 1, 2, \dots, q$,
- $\beta_j \geq 0$ dla $j = 1, 2, \dots, p$.

Uwaga. Model GARCH jest często wykorzystywany do opisu reszt w modelach ARMA.

Właściwości

- Przy spełnieniu odpowiednich założeń model GARCH można sprowadzić do postaci modelu ARCH.
- Model GARCH (p, q) można przedstawić w postaci modelu ARMA (m, p) dla szeregu r_t^2 , gdzie $m = \max(p, q)$.

Stacjonarność

Stacjonarne szeregi czasowe typu GARCH (p, q) spełniają nierówność

$$\sum_{j=1}^q \alpha_j + \sum_{j=1}^p \beta_j < 1.$$

Estymacja

Estymacja modeli klasy GARCH odbywa się z wykorzystaniem metody największej wiarygodności. Maksymalizuje się funkcję log-wiarygodności

$$L_T(\boldsymbol{\theta}) = \sum_{t=1}^T \ell_t(\boldsymbol{\theta}),$$

gdzie

- $\boldsymbol{\theta}$ — wektor wszystkich parametrów modelu,
- $\ell_t(\boldsymbol{\theta})$ — funkcja log-wiarygodności dla obserwacji z chwili czasu t .

Estymacja

Przypadek reszt o rozkładzie normalnym

Niech dane będzie równanie

$$r_t = \sigma_t \epsilon_t,$$

gdzie ϵ_t iid $\mathcal{N}(0, 1)$.

Funkcja log-wiarogodności dla tego równania w chwili czasu t wyraża się wzorem

$$\ell_t(\theta) = \log \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{r_t^2}{2\sigma_t^2}} = -\frac{r_t^2}{2\sigma_t^2} - \log \sigma_t - \log \sqrt{2\pi}.$$

Finalna postać tej funkcji zależy od sposobu zdefiniowania warunkowej wariancji σ_t^2 w danym modelu klasy GARCH.

Ocena jakości modelu

Idea

Idea

Inwestor jest bardziej zainteresowany poznaniem kierunku zmian kursu instrumentu finansowego, niż dokładną wartością tego kursu.

Ocena jakości modelu

Miary zgodności kierunku zmian stopy zwrotu

Niech:

- r_t — rzeczywista wartość stopy zwrotu w chwili czasu t ,
- \hat{r}_t — predykowana wartość stopy zwrotu w chwili czasu t .

Podstawowe miary zgodności kierunku zmian stopy zwrotu to:

$$Q_1 = \frac{\#\{t : r_t \hat{r}_t > 0\}}{\#\{t : r_t \hat{r}_t \neq 0\}},$$
$$Q_2 = \frac{\#\{t : r_{t-1} r_t < 0 \wedge r_t \hat{r}_t > 0\}}{\#\{t : r_{t-1} r_t < 0 \wedge r_t \hat{r}_t \neq 0\}}.$$

Ocena jakości modelu

Modyfikacje miary Q_1

Uwzględniając koszt transakcji w wysokości g możemy miarę Q_1 zmodyfikować w następujący sposób:

$$Q'_1 = \frac{\#\{t : r_t > g \wedge r_t \hat{r}_t > 0\}}{\#\{t : r_t > g \wedge r_t \hat{r}_t \neq 0\}}.$$

W przypadku modeli pozwalających szacować wyłącznie zmienność stopy zwrotu σ_t (np. modele GARCH bez części ARIMA) możemy zastosować następującą modyfikację miary Q_1 :

$$Q_1^x = \frac{\#\{t : r_t \Delta \hat{\sigma}_t < 0\}}{\#\{t : r_t \Delta \hat{\sigma}_t \neq 0\}}.$$

Model GARCH-M

Model GARCH-M

Szereg czasowy y_t jest procesem GARCH-M (p, q) , jeżeli

$$y_t = \lambda + \delta \sigma_t + r_t,$$

$$r_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j r_{t-j}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,$$

gdzie

- ϵ_t iid WN $(0, 1)$,
- $\alpha_0 > 0$,
- $\alpha_j \geq 0$ dla $j = 1, 2, \dots, q$,
- $\beta_j \geq 0$ dla $j = 1, 2, \dots, p$,
- λ, δ stałe.

Właściwości

- Model GARCH-M (p, q) jest rozszerzeniem modelu GARCH (p, q) o dodanie heteroskedastyczności do średniej równania.
- Parametr λ oznacza średnią modelu GARCH
- Parametr σ interpretowany jest jako współczynnik zmienności dla średniej.

Model EGARCH

Model EGARCH

Szereg czasowy r_t jest procesem EGARCH (p, q) , jeżeli

$$\begin{aligned}r_t &= \sigma_t \epsilon_t, \\ \ln \sigma_t^2 &= \alpha_0 + \sum_{j=1}^q \alpha_j r_{t-j}^2 + \sum_{j=1}^p \beta_j \ln \sigma_{t-j}^2,\end{aligned}$$

gdzie

- ϵ_t iid WN $(0, 1)$,
- $\alpha_0 > 0$,
- $\alpha_j \geq 0$ dla $j = 1, 2, \dots, q$,
- $\beta_j \geq 0$ dla $j = 1, 2, \dots, p$.

Model TGARCH

Model TGARCH

Szereg czasowy r_t jest procesem TGARCH (p, q) , jeżeli

$$\begin{aligned}r_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^q \alpha_j r_{t-j}^2 + \sum_{j=1}^q \alpha_j S_{t-j} r_{t-j}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,\end{aligned}$$

gdzie

- ϵ_t iid WN $(0, 1)$,
- $\alpha_0 > 0$,
- $\alpha_j \geq 0$ dla $j = 1, 2, \dots, q$,
- $\beta_j \geq 0$ dla $j = 1, 2, \dots, p$,
- $S_{t-j} = \begin{cases} 1, & \text{dla } r_{t-j} < 0 \\ 0, & \text{dla } r_{t-j} \geq 0 \end{cases}$

Kointegracja

Idea. Chcemy wiedzieć, czy bazując na pewnej grupie niestacjonarnych szeregów czasowych możemy bezpiecznie zbudować model.

Kointegracja

Powiemy, że szeregi czasowe $X_{1,t}, X_{2,t}, \dots, X_{n,t}$ są skointegrowane w stopniu (d, b) , jeżeli dla każdego $i = 1, 2, \dots, n$ zachodzi

$$X_{i,t} \sim I(d)$$

oraz istnieją takie wartości $\beta_1, \beta_2, \dots, \beta_n$, że

$$\beta_1 X_{1,t} + \beta_2 X_{2,t} + \dots + \beta_n X_{n,t} \sim I(d - b).$$

Intuicja. Relacje pomiędzy skointegrowanymi niestacjonarnymi szeregami czasowymi pozostają w długiej perspektywie czasowej niezmiennie.

Testowanie kointegracji

Procedura Engle'a-Grangera.

- 1 Należy zweryfikować czy wszystkie analizowane szeregi czasowe charakteryzuje ten sam stopień integracji.
- 2 Należy zbudować model regresji liniowej wielorakiej w którym jeden z analizowanych szeregów pełni rolę zmiennej objaśnianej, pozostałe natomiast zmiennych objaśniających.
- 3 Należy przetestować stopień integracji reszt wyznaczonego w poprzednim kroku modelu.

Model ECM

Model korekty błędem (ECM)

$$\Delta y_t = \mu + \alpha (y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + \sum_{i=1}^{k-1} \theta_i \Delta y_{t-i} + \sum_{i=0}^{k-1} \gamma_i \Delta x_{t-i} + \epsilon_t$$

Interpretacja:

- $y_{t-1} = \beta_0 + \beta_1 x_{t-1}$ — równanie równowagi długookresowej,
- $y_{t-1} - \beta_0 - \beta_1 x_{t-1}$ — odchylenie od równowagi długookr.,
- α — współczynnik opisujący szybkość dostosowywania się zmiennej objaśnianej do poziomu równowagi długookresowej (w stabilnym modelu $\alpha < 0$).
- θ_i, γ_i — współczynniki opisujące dynamikę krótkookresową.

Stosowalność

Uwaga. Twierdzenie Grangera o reprezentacji gwarantuje nam możliwość zastosowania mechanizmu korekty błędem względem skointegrowanych szeregów czasowych.

Estymacja

- 1 Estymacja parametrów równania równowagi długookresowej

$$y_{t-1} = \beta_0 + \beta_1 x_{t-1}.$$

- 2 Skonstruowanie szeregów czasowych

$$\epsilon_t = y_t - \beta_0 - \beta_1 x_t,$$

$$\Delta x_t = x_t - x_{t-1},$$

$$\Delta y_t = y_t - y_{t-1}.$$

- 3 Estymacja parametrów równania modelu korekty błędem

$$\Delta y_t = \mu + \alpha \epsilon_{t-1} + \sum_{i=1}^{k-1} \theta_i \Delta y_{t-i} + \sum_{i=0}^{k-1} \gamma_i \Delta x_{t-i} + \epsilon_t$$

Pytania?

Dziękuję za uwagę!