AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

Wydział Zarządzania

Projekt EFiD – Projekt 4A

Autorzy:	*Patrycja Piła, Mateusz Strojek, Julia Szutka, Magdalena Wnuk*
Kierunek studiów:	*Informatyka i Ekonometria*
Prowadzący:	*dr hab. Tomasz Wójtowicz*

Kraków, 2025

Polecenie

Część A Przygotować arkusz w Excelu, w którym będzie można za pomocą warunkowej MNW oszacować parametry odpowiedniego modelu ARMA. Przyjąć założenie, że składniki losowe mają rozkład t Studenta. Opisać zastosowane metody estymacji, wyjaśnić w jaki sposób za pomocą tego arkusza dokonać estymacji, wyjaśnić zawartość poszczególnych komórek. Uzupełnić arkusz z projektu 3 tak, by obliczał też wartość prognoz z oszacowanego modelu.

Wstęp

Celem niniejszego zadania jest przygotowanie arkusza kalkulacyjnego w Excelu umożliwiającego estymację parametrów modelu ARMA(2,0) metodą warunkowej największej wiarygodności przy założeniu, że składnik losowy ma rozkład t Studenta. Model ten opisuje zależność wartości zmiennej czasowej od jej dwóch poprzednich opóźnień, co pozwala uchwycić dynamikę i strukturę autokorelacji w danych.

Metoda warunkowej największej wiarygodności polega na maksymalizacji funkcji wiarygodności, przy założeniu znanych początkowych wartości procesu.

Metoda Największej Wiarygodności (reszty z rozkładu normalnego)

Parametry:

- T - liczba obserwacji,
- ϕ_0 - wyraz wolny,
- ϕ_1 - parametr dla pierwszego opóźnienia,
- ϕ_2 - parametr dla drugiego opóźnienia,
- σ^2 – wariancja składnika losowego,
- p – rząd opóźnień.

Formuły w komórkach

Na podstawie szeregu czasowego *Xt* utworzono dwie kolumny z jego opóźnieniami:

- X_{t-1}
- X_{t-2}

Następnie obliczono wartość teoretyczną zmiennej, oznaczoną jako **x_hat**. Została ona wyznaczona ze wzoru:

$$X\_hat = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2}$$

Następnie obliczono reszty modelu:

$$\varepsilon t = Xt - X\_hat$$

Dla każdej obserwacji obliczono składnik funkcji log-wiarygodności w postaci

likelihood sig $= \varepsilon t\text{^}2/(2\sigma\text{^}2)$

Funkcja celu

Funkcja log-wiarygodności jest podana wzorem:

$$\begin{aligned} \log f_{Y_T, Y_{T-1}, \ldots, Y_{p+1} | Y_p, \ldots, Y_1}(y_T, y_{T-1}, \ldots, y_{p+1} | y_p, \ldots, y_1; \boldsymbol{\theta}) \\ = -\frac{T-p}{2}\log(2\pi) - \frac{T-p}{2}\log(\sigma^2) \\ - \sum_{t=p+1}^{T} \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2}. \end{aligned} \qquad [5.3.9]$$

Rysunek 1 Wzór z książki „Time Series Analysis" James D. Hamilton 1994 r.

Po obliczeniu dwóch pomocniczych składników funkcji log-wiarygodności, zostały one **połączone w jedną formułę**, aby uzyskać końcową wartość logarytmu funkcji wiarygodności dla całego modelu.

-(T-p)/2 * log(2pi)	-96,48854599
-(T-p)/2 * log(sigm^2)	400,7562039
Funkcja celu	251,7676039

Za pomocą dodatku Solver przeprowadzono proces maksymalizacji funkcji celu, odpowiadającej logarytmowi funkcji wiarygodności modelu. Optymalizacja została wykonana względem czterech parametrów modelu: $\phi 0, \phi 1, \phi 2,$ σ^2. Dla parametru $\sigma\text{^}2$ wprowadzono ograniczenie dolne równe 0,0001, co zapewnia spełnienie warunku dodatniości wariancji składnika losowego (wariancja musi być większa od zera).

Parametr	Wartość
ϕ_0	-0,003407035
ϕ_1	-0,140986288
ϕ_2	-0,144095741
σ^2	0,000483987

Prognozowanie

Prognozowanie w modelach autoregresyjnych polega na wyznaczaniu przyszłych wartości zmiennej na podstawie jej własnych opóźnień oraz oszacowanych parametrów modelu. W przypadku modelu AR(2) przyjmuje się postać:

$$\hat{Y}_{t+h} = \phi_0 + \phi_1 \hat{Y}_{t+h-1} + \phi_2 \hat{Y}_{t+h-2}$$

Gdzie**:**

- $\hat{Y}_{t+h}$ to prognoza na h okresów do przodu,
- $\hat{Y}_{t+h-1}$, $\hat{Y}_{t+h-2}$ to prognoza na h-1 okresów do przodu i h-2 okresów do przodu,
- ϕ_0, ϕ_1, ϕ_2 to oszacowane parametry modelu,
- t to liczba obserwacji.

Horyzont prognozy	**Wartość**
1	0,0021
2	-0,0002
3	-0,0037

Tabela 1. Wyniki prognoz dla modelu AR(2)

Prognozy wskazują, że w najbliższym okresie oczekiwany jest niewielki wzrost (0,0021), po którym następuje stabilizacja, a następnie lekki spadek wartości zmiennej. Wraz z wydłużeniem horyzontu prognozy wartości maleją, co oznacza, że wpływ wcześniejszych obserwacji słabnie, a proces dąży do równowagi. Zmiany są niewielkie, więc model sugeruje stabilność badanego zjawiska w krótkim okresie.

Metoda Największej Wiarygodności (reszty z rozkładu t- Studenta)

Gęstość rozkładu t- Studenta ma postać:

$$f_{std}(x;v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{x}{2}\right)}\left(1+\frac{x^2}{v}\right)^{-\frac{v+1}{2}}$$

Natomiast z parametrami $v \ i \ \sigma$ można ją zapisać jako:

$$f(\varepsilon_t; v, \sigma) = \frac{1}{\sigma} f_{std}\left(\frac{\varepsilon_t}{\sigma}; v\right)$$

co jest równoznaczne z

$$f(\varepsilon_t; v, \sigma) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{\varepsilon_t^2}{v\sigma^2}\right)^{-\frac{v+1}{2}}$$

Zatem funkcja największej log wiarygodności ma postać:

$$\mathcal{L} = \sum_{t=3}^{T}\left[A - \frac{v+1}{2}\ln\left(1+\frac{\varepsilon_t^2}{v\sigma^2}\right)\right]$$

gdzie $A = \ln\left(\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\right)$ to stała dla każdej obserwacji, a $\varepsilon_t = \ y_t - c - \varphi_1 y_{t-1} - \varphi_2 y_{t-2}$.

Zatem dla punktu t:

$$lnf(\varepsilon_t; v; \sigma) = \ln\left(\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\right) - \frac{v+1}{2}\ln\left(1 + \frac{\varepsilon_t^2}{v\sigma^2}\right)$$

Parametry:

- ϕ_0 - wyraz wolny,
- ϕ_1 - parametr dla pierwszego opóźnienia,
- ϕ_2 - parametr dla drugiego opóźnienia,
- σ – parametr skali
- ν – liczba stopni swobody

Ograniczenia:

Aby model ar(2) był stacjonarny, muszą być spełnione jednocześnie trzy warunki:

- $\phi_1 + \phi_2 < 1$
- $\phi_2 - \phi_1 < 1$
- $-1 < \phi_2 < 1$

Dodatkowo użyto ograniczeń na parametry rozkładu reszt:

- $\sigma > 0$
- $\nu > 2$ (by wariancja była określona)

Ostatecznie po zastosowaniu Solver'a z powyższymi ograniczeniami otrzymano poniższe wartości parametrów:

Parametr	Wartość
ϕ_0	-0,0037
ϕ_1	-0,0847
ϕ_2	-0,1120
σ	0,0149
ν	3,6650

Tabela 2. Oszacowane parametry modelu ARMA przy założeniu rozkładu t Studenta

Parametr $\phi 0$ (wyraz wolny) jest ujemny, co sugeruje, że proces oscyluje wokół wartości nieco poniżej zera. Współczynniki $\phi 1$ i $\phi 2$ również są ujemne, co oznacza, że wzrost wartości zmiennej w poprzednich okresach wpływa negatywnie na jej obecną wartość.

Odchylenie standardowe składnika losowego ($\sigma=0{,}0149$) wskazuje na umiarkowaną zmienność procesu. Liczba stopni swobody rozkładu t ($\nu=3{,}665$) potwierdza obecność „grubszych ogonów" niż w rozkładzie normalnym, co oznacza, że dane mogą zawierać obserwacje odstające.