Badanie stacjonarności, modele o rozłożonych opóźnieniach i modele autoregresyjne o rozłożonych opóźnieniach

Stanisław Cichocki
Natalia Nehrebecka

Wykład 5

Plan wykładu

- 1. Badanie stacjonarności:
  - Test Dickey-Fullera (DF)
  - Rozszerzony test Dickey-Fullera (ADF)
  - Test KPSS
- 2. Modele o rozłożonych opóźnieniach (DL)
- 3. Modele autoregresyjne o rozłożonych opóźnieniach (ADL)

Plan wykładu

- 1. Badanie stacjonarności:
  - Test Dickey-Fullera (DF)
  - Rozszerzony test Dickey-Fullera (ADF)
  - Test KPSS
- 2. Modele o rozłożonych opóźnieniach (DL)
- 3. Modele autoregresyjne o rozłożonych opóźnieniach (ADL)

Test Dickey-Fullera (DF)

- Równanie regresji:

$$\Delta y_t = \alpha_1 + \rho y_{t-1} + \varepsilon_t$$

$H_0 : \rho = 0, \alpha_1 = 0$ - błądzenie przypadkowe

$H_1 : \rho \in (-2,0)$ - model AR(1) ze stałą

Test Dickey-Fullera (DF)

- Równanie regresji:

$$\Delta y_t = \alpha_1 + \alpha_2 t + \rho y_{t-1} + \varepsilon_t$$

$H_0 : \rho = 0, \alpha_2 = 0$ - błądzenie przypadkowe z dryfem

$H_1 : \rho \in (-2,0)$ - model AR(1) z trendem liniowym

Rozszerzony test Dickey-Fullera (ADF)

- Często reszty z regresji:

$$\Delta y_t = \rho y_{t-1} + \varepsilon_t$$

wykazują silną autokorelację, co jest problemem.

- Rozszerzony test Dickey-Fullera (test ADF) różni się od standardowego testu DF rozszerzeniem regresji o dodatkowe elementy, których celem jest eliminacja autokorelacji reszt.

Rozszerzony test Dickey-Fullera (ADF)

- Przeprowadzamy regresję:

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta y_{t-i} + \varepsilon_t$$

gdzie: $\gamma_i \Delta y_{t-i}$ - rozszerzenie

- Liczbę opóźnień k dobieramy tak aby z reszt wyeliminować autokorelację.

Rozszerzony test Dickey-Fullera (ADF)

- Statystyka testowa dla testu ADF : statystyka t policzona dla oszacowania parametru przy y_{t-1}

- Dla dużych prób tablice wartości krytycznych dla testu ADF są takie same jak w teście DF.

- Dla małych prób, małopróbkowe wartości krytyczne testu DF są jedynie aproksymacją prawdziwych wartości krytycznych testu ADF.

Rozszerzony test Dickey-Fullera (ADF)

- Test ADF przeprowadzamy w następujący sposób:

  1. regresja Δy_t na y_{t-1}
  2. badamy występowanie autokorelacji za pomocą testu Breuscha-Godfreya (nie Durbina-Watsona):

     a) jeśli autokorelacja nie występuje to porównujemy statystykę t dla y_{t-1} z wartościami krytycznymi testu ADF:

        i. wartość statystyki jest mniejsza od wartości krytycznej - odrzucamy H_0 o niestacjonarności y_t i przyjmujemy H_1 o stacjonarności y_t ;

Rozszerzony test Dickey-Fullera (ADF)

ii. wartość statystyki jest większa od wartości krytycznej – nie ma podstaw do odrzucenia H_0

b) jeśli autokorelacja występuje to dodajemy stopniowo rozszerzenia do momentu aż się jej pozbędziemy a następnie porównujemy statystykę t dla y_{t-1} z wartościami krytycznymi testu ADF

i. wartość statystyki jest mniejsza od wartości krytycznej - odrzucamy H_0 o niestacjonarności y_t i przyjmujemy H_1 o stacjonarności y_t ;

Rozszerzony test Dickey-Fullera (ADF)

ii. wartość statystyki jest większa od wartości krytycznej – nie ma podstaw do odrzucenia H_0

Rozszerzony test Dickey-Fullera (ADF)

- W przypadku błądzenia przypadkowego:

```
      Source |       SS       df       MS              Number of obs =     999
-------------+------------------------------           F(  1,   998) =    0.08
       Model |  .077155196     1  .077155196           Prob > F      =  0.7815
    Residual |  1000.45476   998  1.00245968           R-squared     =  0.0001
-------------+------------------------------           Adj R-squared = -0.0009
       Total |  1000.53191   999  1.00153345           Root MSE      =  1.0012

------------------------------------------------------------------------------
        D.rw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          rw |
         L1. |  -.0005213    .001879    -0.28   0.782    -.0042086     .003166
------------------------------------------------------------------------------
```

Rozszerzony test Dickey-Fullera (ADF)

- W przypadku błądzenia przypadkowego:

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	1.084	1	0.2977
2	1.477	2	0.4779
3	2.629	3	0.4524
4	2.639	4	0.6199

H0: no serial correlation

Test KPSS

- Test KPSS (Kwiatkowski, Philips, Schmidt, Shin) testuje hipotezę zerową o stacjonarności zmiennej.

- Oparty jest na następującym modelu:

$$y_t = \delta + \zeta_t + \varepsilon_t$$

$$\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

$$\zeta_t = \zeta_{t-1} + u_t$$

$$u_t \sim IID(0, \sigma_u^2)$$

Test KPSS

$H_0 : \sigma_u^2 = 0$ - zmienna y_t jest stacjonarna

$H_1 : \sigma_u^2 > 0$ - zmienna y_t jest niestacjonarna

- Hipotezę zerową odrzucamy gdy statystka testowa > wartości krytycznej.

- Statystyka testowa dla testu KPSS zawsze >0.

Test KPSS

- Test KPSS dla białego szumu:

```
Critical values for H0: x is level stationary

10%: 0.347  5% : 0.463  2.5%: 0.574  1% : 0.739

Lag order    Test statistic
    0            .0629
    1            .0656
    2             .064
    3            .0623
    4            .0607
    5            .0594
    6            .0582
    7            .0582
    8            .0584
    9             .059
   10            .0596
```

Test KPSS

- Test KPSS dla białego szumu:

```
Critical values for H0: x is level stationary

10%: 0.347  5% : 0.463  2.5%: 0.574  1% : 0.739

Lag order    Test statistic
    0            .0629
    1            .0656
    2             .064
    3            .0623
    4            .0607
    5            .0594
    6            .0582
    7            .0582
    8            .0584
    9             .059
   10            .0596
```

Test KPSS

- Test KPSS dla błądzenia przypadkowego

```
Critical values for H0: rw is level stationary

10%: 0.347  5% : 0.463  2.5%: 0.574  1% : 0.739

Lag order    Test statistic
    0            55.9
    1              28
    2            18.7
    3            14.1
    4            11.3
    5            9.42
    6            8.09
    7            7.09
    8            6.31
    9            5.69
   10            5.18
```

Test KPSS

- Test KPSS dla błądzenia przypadkowego

```
Critical values for H0: rw is level stationary

10%: 0.347  5% : 0.463  2.5%: 0.574  1% : 0.739

Lag order    Test statistic
    0             55.9
    1               28
    2             18.7
    3             14.1
    4             11.3
    5             9.42
    6             8.09
    7             7.09
    8             6.31
    9             5.69
   10             5.18
```

Plan wykładu

- 1. Badanie stacjonarności:
  - Test Dickey-Fullera (DF)
  - Rozszerzony test Dickey-Fullera (ADF)
  - Test KPSS
- 2. Modele o rozłożonych opóźnieniach (DL)
- 3. Modele autoregresyjne o rozłożonych opóźnieniach (ADL)

Modele o rozłożonych opóźnieniach

- W modelach o rozłożonych opóźnieniach (Distributed Lags) wpływ zmiennych niezależnych na zmienną zależną jest rozłożony w czasie.

$$y_t = \mu + \beta_0 x_t + \cdots + \beta_p x_{t-p} + \varepsilon_t$$

Jeśli x_t jest nielosowe i spełnione są odpowiednie założenia dotyczące składnika losowego (jakie?) to model spełnia założenia KMRL.

Modele o rozłożonych opóźnieniach

- Interpretacja parametrów: β_p - zmiana y_t jaka nastąpi jeśli zmieni się x sprzed p okresów a dla wszystkich pozostałych okresów x pozostanie niezmienione.

- Taka interpretacja jest niepraktyczna stąd stosuje się różne mnożniki.

Modele o rozłożonych opóźnieniach

- Mnożnik bezpośredni (impact multiplier): mierzy wielkość oczekiwanej, natychmiastowej reakcji y_t na zmianę x_t o Δx.

$$E(y_t + \Delta y) = \mu + \beta_0(x_t + \Delta x) + \cdots + \beta_p x_{t-p}$$
$$= \mu + \beta_0 x_t + \cdots + \beta_p x_{t-p} + \beta_0 \Delta x$$
$$= E(y_t) + \beta_0 \Delta x$$

$$\beta_0$$

Modele o rozłożonych opóźnieniach

- Mnożnik skumulowany: mierzy wielkość oczekiwanej reakcji y_t na zmianę x o Δx w kolejnych τ okresach.

$$
\begin{aligned}
&E(y_t + \Delta y) \\
&= \mu + \beta_0(x_t + \Delta x) + \cdots + \beta_\tau(x_{t-\tau} + \Delta x) \\
&+ \beta_{\tau+1}x_{t-\tau+1} + \cdots + \beta_p x_{t-p} \\
&= E(y_t) + \left(\sum_{i=0}^{\tau} \beta_i\right) \Delta x
\end{aligned}
$$

$$\beta_\tau = \sum^{\tau} \beta_i$$

Modele o rozłożonych opóźnieniach

- Mnożnik długookresowy (long-run multiplier): mierzy wielkość oczekiwanej reakcji y_t na zmianę wszystkich przeszłych wartości x o Δx.

- Można go policzyć jako mnożnik skumulowany dla $\tau \to \infty$.

$$\sum_{s=0}^{\infty} \beta_s = \beta$$

Modele o rozłożonych opóźnieniach

- Średnie opóźnienie reakcji y_t na zmiany x.

$$\bar{w} = \sum_{s=1}^{\infty} s \frac{\beta_s}{\beta}$$

Należy pamiętać, że dla modelu w którym pojawia się więcej niż jedna zmienna niezależna wielkości mnożników i średnich opóźnień liczone są osobno dla każdej z tych zmiennych.

Modele o rozłożonych opóźnieniach

- Przykład:

$$konsumpcja_t = \mu + 0{,}4doch\acute{o}d_t + 0{,}3doch\acute{o}d_{t-1} + 0{,}2doch\acute{o}d_{t-2}$$

Mnożnik bezpośredni: $\beta_0 = 0{,}4$

Mnożnik długookresowy: $\beta = 0{,}4 + 0{,}3 + 0{,}2 = 0{,}9$

Średnie opóźnienie: $\bar{w} = 1 * \frac{0{,}4}{0{,}9} + 2 * \frac{0{,}3}{0{,}9} + 3 * \frac{0{,}2}{0{,}9} = 1{,}778$

Plan wykładu

- 1. Badanie stacjonarności:
  - Test Dickey-Fullera (DF)
  - Rozszerzony test Dickey-Fullera (ADF)
  - Test KPSS
- 2. Modele o rozłożonych opóźnieniach (DL)
- 3. Modele autoregresyjne o rozłożonych opóźnieniach (ADL)

Modele autoregresyjne o rozłożonych opóźnieniach

- W modelach autoregresyjnych o rozłożonych opóźnieniach (Autoregressive Distributed Lags) występują dodatkowo opóźnione zmienne zależne.

$$y_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \mu + \beta_0 x_t + \cdots + \beta_s x_{t-s} + \varepsilon_t$$

Modele autoregresyjne o rozłożonych opóźnieniach

- Stan równowagi długookresowej (steady state): jest to stan, w którym wartość oczekiwana zmiennej zależnej pozostaje stała w czasie, o ile tylko nie zmieniają się zmienne niezależne.

$$y^* = E(y_t) = E(y_{t-1}) = \cdots = E(y_{t-p})$$
$$x^* = x_t = x_{t-1} = \cdots = x_{t-s}$$

Modele autoregresyjne o rozłożonych opóźnieniach

- Stan równowagi długookresowej:

$$(1-\alpha_1-\cdots-\alpha_p)y^* = \mu + x^*\beta_0 + x^*\beta_1 + \cdots + x^*\beta_s$$

oznaczając: $\mu^* = \dfrac{\mu}{1-\alpha_1-\cdots-\alpha_p}$ $\quad \beta = \dfrac{\beta_0+\beta_1+\cdots\beta_s}{1-\alpha_1-\cdots-\alpha_p}$

otrzymujemy: $y^* = \mu^* + x^*\beta$

Modele autoregresyjne o rozłożonych opóźnieniach

- Mnożnik bezpośredni: mierzy wielkość oczekiwanej, natychmiastowej reakcji y_t na zmianę x_t o Δx.

$$E(y_t + \Delta y) = E(y_t) + \beta_0 \Delta x$$

$$\beta_0$$

Modele autoregresyjne o rozłożonych opóźnieniach

- Mnożnik długookresowy: mierzy wielkość oczekiwanej reakcji y_t na zmianę wszystkich przeszłych wartości x o Δx.

$$y^* = \mu^* + x^*\beta$$

$$y^* + E(\Delta y) = \mu^* + (x^* + \Delta x)\beta$$

$$\beta = \frac{\beta_0 + \beta_1 + \cdots \beta_s}{1 - \alpha_1 - \cdots - \alpha_p}$$

Modele autoregresyjne o rozłożonych opóźnieniach

- Przykład:

$$\begin{aligned}
logkonsumpcja_t \\
&= 0{,}69 logkonsumpcja_{t-1} \\
&+ 0{,}12 logkonsumpcja_{t-2} - 0{,}4 \\
&+ 0{,}76 logdochód_t - 0{,}65 logdochód_{t-1} \\
&- 0{,}074 logdochód_{t-2}
\end{aligned}$$

Mnożnik bezpośredni: $\beta_0 = 0{,}76$

Mnożnik długookresowy: $\beta = \frac{0{,}76-0{,}65-0{,}074}{1-0{,}69-0{,}12}$ =0,19

Modele autoregresyjne o rozłożonych opóźnieniach

- W przypadku gdy w modelu występują opóźnione zmienne zależne nie można utrzymać założenia, że zmienne niezależne są nielosowe. Dlaczego?

- W przypadku gdy zmienne niezależne są losowe można dowieść zgodność estymatora MNK tylko wtedy, gdy $Cov(x_t, \varepsilon_t) = 0$.

- Powyższe założenie nie jest spełnione w modelach ADL, jeśli wystąpi autokorelacja składnika losowego.

Modele autoregresyjne o rozłożonych opóźnieniach

- Stąd powinniśmy dążyć do eliminacji autokorelacji w modelach ADL poprzez dodawania opóźnionych zmiennych zależnych do modelu lub poprzez modelowanie sposobu w jaki są skorelowane błędy losowe.

- Do wykrywania autokorelacji w modelach ADL stosujemy test Breuscha-Godfreya a nie test Durbina-Watsona. Dlaczego?

Dziękuję za uwagę