Kointegracja

ASC 2023

Piotr Żoch

Plan

- Kointegracja
- Testowanie kointegracji
- Model korekty błędem (ECM)

Regresja pozorna - przypomnienie

Regresja pozorna

- Modelowanie zależności pomiędzy *niepowiązanymi* zmiennymi niestacjonarnymi może skutkować *regresją pozorną*.

$$
\begin{aligned}
x_t &= x_{t-1} + \epsilon_t \\
y_t &= y_{t-1} + \nu_t \\
\epsilon_t &\sim iid\left(0, \sigma_\epsilon^2\right) \\
\nu_t &\sim iid\left(0, \sigma_\nu^2\right)
\end{aligned}
$$

- W przykładzie powyżej x_t i y_t są niezależne, ale w regresji

$$y_t = \beta x_t + \varepsilon_t$$

statystyka t często sugeruje odrzucenie $H_0 : \beta = 0$.

Regresja pozorna

- Typowe postępowanie: zróżnicowanie zmiennych i regresja na różnicach

$$\Delta y_t = \beta \Delta x_t + \varepsilon_t$$

- Różnicując zmienne tracimy część informacji.

Kointegracja

Kointegracja

- Teoria często sugeruje, że między pewnymi zmiennymi powinna być jakaś długookresowa zależność. Przykłady:
    - Podaż pieniądza i inflacja;
    - Inwestycje i konsumpcja;
    - ...

Kointegracja

- Zmienne mogą być niestacjonarne, ale mieć *wspólny* trend stochastyczny.

$$
\begin{aligned}
x_t &= z_t + \epsilon_t \\
y_t &= z_t + \nu_t \\
z_t &= z_{t-1} + \zeta_t \\
\epsilon_t &\sim iid\left(0, \sigma_\epsilon^2\right) \\
\nu_t &\sim iid\left(0, \sigma_\nu^2\right) \\
\zeta_t &\sim iid\left(0, \sigma_\zeta^2\right)
\end{aligned}
$$

- W przykładzie powyżej z_t jest wspólne dla (x_t, y_t), lecz przejściowo x_t i y_t mogą oddalić się od z_t (za sprawą ϵ_t, ν_t).

Kointegracja

- Jeśli będziemy badać związek między Δx_t a Δy_t to stracimy (*cenną!*) informację o *długookresowej* zależności między x_t a y_t...
- W naszym przykładzie

$$\begin{aligned} x_t - y_t &= \epsilon_t - \nu_t \\ &= \xi_t \sim IID\left(0, \sigma_\xi^2\right) \end{aligned}$$

czyli $(x_t - y_t)$ jest białym szumem, mimo że $x_t \sim I(1)$ i $y_t \sim I(1)$.

Kointegracja

- Niech $(x_{1t}, x_{2t}, ..., x_{nt})$ będą zmiennymi zintegrowanymi stopnia d.
- Jeśli istnieje wektor $(\beta_1, \beta_2, ..., \beta_n)$, że:

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

gdzie $b > 0$ to mówimy, że

$$(x_{1t}, x_{2t}, ..., x_{nt})$$

są **skointergowane** stopnia (d, b) a wektor β nazywamy **wektorem kointegrującym**.

Kointegracja

- W naszym przykładzie $x_t \sim I(1)$ oraz $y_t \sim I(1)$

$$\begin{aligned} x_t - y_t &= \epsilon_t - \nu_t \\ &= \xi_t \sim IID\left(0, \sigma_\xi^2\right) \end{aligned}$$

czyli

$$x_t - y_t \sim I(0)$$

a wektor kointegrujący to $(1, -1)$.

Kointegracja

- Kointegracja dotyczy *kombinacji liniowej*. Jeśli istnieje wektor kointegrujący, to istnieje ich wiele (jeśli β jest takim wektorem to $\lambda\beta$ dla $\lambda \neq 0$ też nim jest), więc zwykle normalizujemy $\lambda = 1/\beta_1$.
- Stopień integracji musi być *taki sam* dla wszystkich zmiennych (procesów stochastycznych).
- Jeśli jest n zmiennych, to może być do $n-1$ liniowo niezależnych wektorów kointegrujących (ich liczba to *rząd kointegracji*).

Super-zgodność

- Jeśli (x_t, y_t) są skointegrowane, to estymator MNK β $(\hat{\beta})$ w regresji

$$y_t = \alpha + \beta x_t + u_t$$

  będzie *super-zgodny:* wariancja $\hat{\beta}$ będzie zbiegać szybciej do 0, niż w standardowym przypadku.
- Uwaga: $\hat{\beta}$ ma wtedy niestandardowy rozkład.

Kointegracja

- Jeśli zmienne są skointegrowane, to reszty w regresji

$$y_t = \alpha + \beta x_t + u_t$$

  będą stacjonarne.
- Interpretacja: jeśli z jakiegoś powodu y_t oddali się od $\alpha + \beta x_t$, to z czasem wróci. Długookresowa relacja między y_t a x_t.

Kointegracja

- Co się stanie, jeśli zmienne są skointegrowane, ale tego nie uwzględnimy? Powróćmy do przykładu gdzie

$$x_t - y_t = \epsilon_t - \nu_t$$

czyli

$$\Delta x_t = \Delta\zeta_t + \Delta\epsilon_t$$

$$\Delta y_t = \Delta\zeta_t + \Delta\nu_t$$

$$\begin{aligned}\hat{\beta} &= \frac{Cov\left(\Delta\zeta_t + \Delta\epsilon_t, \Delta\zeta_t + \Delta\nu_t\right)}{Var\left(\Delta\zeta_t + \Delta\nu_t\right)} \\ &= \frac{Var\Delta\zeta_t}{Var\Delta\zeta_t + Var\Delta\nu_t} \neq 1\end{aligned}$$

- Obciążenie.

Testowanie kointegracji

- Krok 1: sprawdzić czy zmienne są zintegrowane tego samego stopnia.
  - Jeśli są stacjonarne, nie ma potrzeby rozważania kointegracji.
  - Jeśli są zintegrowane, ale innego stopnia, nie są wtedy skointegrowane.
- Krok 2: estymacja

$$y_t = \alpha + \beta x_t + u_t$$

  i analiza stacjonarności reszt, $\hat{u}_t$.

Test Engle-Granger

- Jak zbadać stacjonarność reszt? Regresja

$$\Delta\hat{u}_t = \rho\hat{u}_{t-1} + e_t$$

(czyli taka sama logika jak w przypadku testu DF).

$$H_0 : \rho = 0$$
$$H_1 : \rho < 0$$

Test Engle-Granger

- Uwaga: nie można użyć wartości krytycznych z testu (A)DF!
- Przyczyna: $\hat{u}_t$ zostały skonstruowane z regresji liniowej, nie znamy prawdziwych (α, β) - jedynie ich oszacowania.
- Wartości krytyczne zależą między innymi od tego, czy y_t i x_t to błądzenie losowe z dryfem, czy z trendem i czy w regresji uwzględniliśmy stałą (α)...
- Jeśli występuje autokorelacja reszt $\hat{e}_t$, uwzględnić opóźnienia $\Delta \hat{u}_t$:

$$\Delta \hat{u}_t = \rho \hat{u}_{t-1} + \sum \Delta \hat{u}_{t-i} + e_t$$

Testowanie kointegracji

- Badanie kointegracji między więcej niż 2 zmiennymi nieco trudniejsze.
    - powtarzanie testu EG, wpierw między wszystkimi zmiennymi, potem między grupami $n-1$ zmiennych i tak dalej.
    - test Johansena - sprawdza rząd kointegracji.

Model korekty błędem (ECM)

Model korekty błędem

- Kointegracja jest sposobem opisu równowagi długookresowej.
- **Twierdzenie Grangera o reprezentacji:** jeśli zmienne losowe (x_t, y_t) są $I(1)$ oraz skointegrowane (z wektorem kointegrującym $(1, -\beta)$) to zależność między nimi można przedstawić za pomocą model korekty błędem (ECM):

$$\Delta y_t = \kappa + \sum_{i=0}^{K-1} \gamma_i \Delta x_{t-i} + \sum_{i=0}^{L-1} \theta_i \Delta y_{t-1} + \boldsymbol{\delta}\left(\mathbf{y_{t-1}} - \boldsymbol{\alpha} - \boldsymbol{\beta}\mathbf{x_{t-1}}\right) + \varepsilon_t$$

gdzie $\varepsilon_t \sim I(0)$ a $\delta \in (-1, 0)$

Model korekty błędem

- Jeżeli zmienne są skointegrowane, to można je opisać modelem korekty błędem.
- Jeżeli zmienne są opisane modelem korekty błędem, to są skointegrowane.

Model korekty błędem

- Jeśli zmienne losowe (x_t, y_t) są $I(1)$ oraz skointegrowane (z wektorem kointegrującym $(1, -\beta)$) to

$$y_t - \alpha - \beta x_t \equiv e_t \sim I(0)$$

  a e_t to odchylenie od długookresowej równowagi.
- β to elastyczność *długookresowa*.

Model korekty błędem

- Model korekty błędem to:

$$\Delta y_t = \kappa + \boldsymbol{\delta e_{t-1}} + \cdots + \varepsilon_t$$

- Parametr $\delta \in (-1, 0)$ to tempo powrotu do równowagi dlugookresowej.
  - okres połowicznego wygaśnięcia to

$$-\frac{\ln 2}{\ln(1+\delta)}$$

- Model uwzględnia informację o tendencji do powrotu do długookresowej równowagi:
  - jeśli $y_{t-1} > \alpha + \beta x_{t-1}$ to $e_{t-1} > 0$ a zatem można spodziewać się, że $\Delta y_t < 0$.