Modele wektorowej autoregresji

ASC 2023

Piotr Żoch

Plan

- Model VAR
- Budowa i analiza modelu VAR
- Prognozowanie za pomocą modelu VAR
- Modele SVAR

Modele VAR

Model VAR

- VAR: **v**ector **a**uto**r**egression
- uwzględnia powiązania między zmiennymi ekonomicznymi
- pomocny w opisie dynamicznych właściwości zmiennej wielowymiarowej
- przy dodatkowych założeniach - model ekonomiczny (structural VAR - SVAR)

Model VAR

- Modele VAR powstały w późnych latach 70. w odpowiedzi na problemy dużych modeli strukturalnych:
    - podział na zmienne endo- i egzogeniczne a priori
    - wybór struktury dynamicznej modelu a priori
    - niska jakość prognoz
    - problemy z identyfikacją

Model VAR

- Sims (1980), modele VAR przydatne w:
    - prognozowaniu szeregów czasowych
    - projektowaniu i ewaluacji modeli ekonomicznych
    - ocenie efektów alternatywnych polityk (jeśli nie różnią się zbytnio od obecnych)

Model VAR

- Niech y_t będzie wektorem $y_t = (y_{1,t}, y_{2,t}, ..., y_{n,t})'$. Model VAR to

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \epsilon_t$$

  gdzie $\epsilon_t \sim N(0, \Sigma)$.
- A_0 to wektor $n \times 1$, , $A_1, ..., A_p$ oraz Σ to macierze o wymiarach $n \times n$.
- Można też dodać zmienne deterministyczne, np. trend, zmienne 0/1, zmienne sezonowe i tym podobne.

Model VAR

- Model VAR można też przedstawić za pomocą wielomianu opóźnień jako

$$(I - A_1 L - \ldots - A_p L^p)\, y_t = A_0 + \epsilon_t$$

czyli

$$A(L)\, y_t = A_0 + \epsilon_t.$$

Właściwości modeli VAR

- Brak podziału a priori na zmienne endo- i egzogeniczne
- Ateoretyczność (teoria służy do wybrania zmiennych, nie do modelowania zależności między nimi)
- Zwykle bardzo dobre prognozy w krótkim okresie
- Parametry modelu *nie* podlegają interpretacji, zamiast tego interesuje nas np. funkcja reakcji na impuls

Budowa modelu VAR

- Dwie decyzje przy budowie modelu VAR:
  - jakie zmienne uwzględnić (+ zmienne deterministyczne)
  - jaki rząd opóźnień wybrać

Dobór rzędu opóźnień

- Kryteria:
    - kryteria informacyjne (AIC, BIC, HQIC)
        - dopasowanie do danych vs. liczba oszacowanych parametrów
    - brak autokorelacji składnika losowego
    - test istotności opóźnień

VAR - stacjonarność

- W przypadku modelu AR(p) mieliśmy

$$\left(1 - a_1 L - a_2 L^2 - \ldots - a_p L^p\right) y_t = a_0 + \epsilon_t$$
$$A\left(L\right) y_t = a_0 + \epsilon_t$$

- Model AR(p) jest **stacjonarny** tylko dla wartości parametrów $a_1, \ldots, a_p$, dla których pierwiastki równania $A\left(x\right)$ spełniają $|x_p| > 1$ dla każdego p
- Analogicznie jest w przypadku modelu VAR(p)

VAR - stacjonarność

- Jeśli $\lim_{k\to\infty} \frac{\partial y_{t+k}}{\partial \epsilon_t} = 0$ to model VAR jest *stacjonarny*
  - wpływ szoku w okresie t wygasa w przyszłości
    - istnieje poziom μ, do którego proces powraca
    - tempo powrotu dane przez pierwiastki równania

$$|A(z)| = 0$$

    - jeśli pierwiastki są poza kołem jednostkowym - stacjonarność

VAR - stacjonarność

- Przykład: VAR(1)

$$y_t = A_1 y_{t-1} + \epsilon_t$$

$$\underbrace{(I - A_1 L)}_{=A} y_t = \epsilon_t$$

czyli

$$A = \begin{bmatrix} 1 - a_{11}L & -a_{12}L \\ -a_{21}L & 1 - a_{22}L \end{bmatrix}$$

i

$$|A(z)| = (1 - a_{11}z)(1 - a_{22}z) - a_{11}a_{21}z^2$$

VAR - stacjonarność

- Przykład: VAR(1). Niech $a_{11} = a_{22} = \rho$ a $a_{12} = a_{21} = 0$.

  $$|A(z)| = (1 - \rho z)(1 - \rho z)$$

  więc $z = \rho^{-1}$ a zatem pierwiastki będą poza kołem jednostkowym dla $-1 < \rho < 1$.

VAR - stacjonarność

- Co w przypadku VAR(p), gdzie $p > 1$?
- Każdy VAR(p) można przekształcić do VAR(1) *(postać kanoniczna)*
- Przykład: AR(2)

$$y_{1,t} = \rho_1 y_{1,t-1} + \rho_2 y_{1,t-2} + \epsilon_t$$

Zdefiniujmy $y_{2,t} = y_{1,t-1}$. Mamy układ równań

$$y_{1,t} = \rho_1 y_{1,t-1} + \rho_2 y_{2,t-1} + \epsilon_t$$

$$y_{2,t} = y_{1,t-1}$$

czyli

$$y_t = \underbrace{\begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix}}_{A_1} y_{t-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \epsilon_t$$

VAR - stacjonarność

- Dla VAR(1)

$$|A(z)| = |I - A_1 z|$$

więc jeśli mamy z, że $|I - A_1 z| = 0$ to

$$|I \frac{1}{z} - A_1| = 0$$

czyli

$$|A_1 - I \underbrace{\lambda}_{=\frac{1}{z}}| = 0$$

VAR - stacjonarność

- VAR(1) jest stacjonarny, jeśli wszystkie pierwiastki λ równania

$$|A_1 - I\lambda| = 0$$

  leżą w kole jednostkowym.
- λ to wartości własne macierzy A_1
- Uwaga: łatwo się pomylić czy pierwiastki/wartości własne mają być *w* czy *poza* kołem jednostkowym (podobnie jak w przypadku modeli AR)

Estymacja parametrów modelu VAR

- Estymacja: można zastosować metodę najmniejszych kwadratów (KMNK/OLS) dla każdego równania osobno
- Co zrobić jeśli poszczególne zmienne są niestacjonarne?
    - uwzględnić informację o kointegracji (jeśli jest) - model VECM
    - pierwsze różnice zamiast poziomów
    - nic (Sims et al. 1990) - interesują nas zależności między zmiennymi, a nie parametry
- Najczęstsza (i najbezpieczniejsza praktyka): dodać trend deterministyczny

Testy na autokorelację reszt

- Autokorelacja:
  - estymator MNK jest obciążony
  - oznaka złej specyfikacji modelu VAR (zbyt niski rząd opóźnień)
- Podstawowe testy:
  - Ljung-Box
  - Breusch-Godfrey

Testy restrykcji

- Test ilorazu wiarygodności: test istotności kolejnych opóźnień wszystkich zmiennych w poszczególnych równaniach.
  - H_0 : zasadność nałożonych restrykcji
- Idea: porównać funkcję wiarygodności modelu z restrykcjami (mniejsza liczba opóźnień) i bez (większa liczba opóźnień).

Prognozowanie z modelem VAR

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + ... A_p y_{t-p} + \epsilon_t$$

czyli

$$E_T\left[y_{T+1}\right] = A_0 + A_1 y_T + A_2 y_{T-1} + ... A_p y_{T-p+1}$$
$$E_T\left[y_{T+2}\right] = A_0 + A_1 E_T\left[y_{T+1}\right] + A_2 y_T + ... A_p y_{T-p+2}$$

i tak dalej.

Prognozowanie z modelem VAR

- Stacjonarny proces VAR można zapisać w postaci wektorowej średniej ruchomej (VMA)

$$y_t = \mu + \sum_{i=0}^{\infty} \Theta_i \epsilon_{t-i}$$

czyli

$$y_{T+1} - E_T\left[y_{T+1}\right] = \epsilon_{T+1}$$
$$y_{T+2} - E_T\left[y_{T+2}\right] = \epsilon_{T+2} + \Theta_1 \epsilon_{T+1}$$

i tak dalej.

- Wariancja prognozy na k okresów

$$Var\left[y_{T+k} - E_T\left[y_{T-k}\right]\right] = \sum_{i=0}^{k-1} \Theta_i \Sigma \Theta_i'$$

co pozwala nam zbudować przedziały ufności.

Modele SVAR

VAR i SVAR

- Model VAR

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \epsilon_t$$
$$\epsilon_t \sim N(0, \Sigma)$$

- Model SVAR

$$Ay_t = C_0 + C_1 y_{t-1} + C_2 y_{t-2} + \ldots + C_p y_{t-p} + B\eta_t$$
$$\eta_t \sim N(0, I)$$

- Różnica: składniki losowe w SVAR są względem siebie niezależne. Interpretacja: *szoki strukturalne*.

Model SVAR

- Model SVAR

$$
\begin{aligned}
Ay_t &= C_0 + C_1 y_{t-1} + C_2 y_{t-2} + \ldots + C_p y_{t-p} + B\eta_t \\
\eta_t &\sim N(0, I)
\end{aligned}
$$

- Macierze A i B określają jednoczesne zależności między zmiennymi endogenicznymi, macierze C_p dynamiczne właściwości modeli.

Postać strukturalna i zredukowana

- Rozważmy prosty model SVAR

$$
\begin{aligned}
y_t &= C y_{t-1} + B\eta_t \\
\eta_t &\sim N(0, I)
\end{aligned}
$$

  powyższą postać nazywamy *postacią strukturalną*.
- Problem: nie obserwujemy B i η_t oddzielnie (obserwujemy jedynie y_t). *Postać zredukowana*:

$$
\begin{aligned}
y_t &= C y_{t-1} + \epsilon_t \\
\epsilon_t &\sim N\left(0, \underbrace{BB'}_{=\Sigma}\right)
\end{aligned}
$$

Postać strukturalna i zredukowana

- W ogólnym przypadku SVAR (postać strukturalna)

$$Ay_t = C_0 + C_1 y_{t-1} + C_2 y_{t-2} + \ldots + C_p y_{t-p} + B\eta_t$$
$$\eta_t \sim N(0, I)$$

- Postać zredukowana

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \epsilon_t$$
$$\epsilon_t \sim N(0, \Sigma)$$

gdzie $A_0 = A^{-1}C_0$... oraz $\epsilon_t = A^{-1}B\eta_t$ i $\Sigma = A^{-1}B\left(A^{-1}B\right)'$.

Model SVAR

- Parametry modelu SVAR otrzymuje się w dwóch krokach:
    1. estymacja parametrów modelu VAR (postać zredukowana)
    2. strukturalizacja/identyfikacja ("przejście" z ϵ_t do η_t)

Funkcja reakcji na impuls

- Model SVAR pozwala analizować dynamiczną reakcję na szoki strukturalne
  - *"jak zmieni się inflacja za 3 kwartały, jeśli stopa procentowa zostanie teraz niespodziewanie podniesiona o 25pb?"*
- Reakcja określana przez funkcję reakcji na impuls (impulse response function - IRF)

$$IRF_{k(ij)} = \frac{\partial y_{i,t+k}}{\partial \eta_{jt}}$$

Funkcja reakcji na impuls

- IRF można otrzymać przez przekształcenie modelu do postaci SVMA(∞):

$$y_t = \mu + \sum_{k=0}^{\infty} \Psi_k \eta_{t-k}$$

gdzie $\Psi_k = IRF_k$.

Funkcja reakcji na impuls

- Aby otrzymać IRF:
  - Zapisać model SVAR w postaci *zredukowanej* (VAR):

$$y_t = A^{-1}C_0 + A^{-1}C_1 y_{t-1} + A^{-1}C_2 y_{t-2} + \ldots + A^{-1}C_p y_{t-p} + A^{-1}B\eta_t$$

    czyli $\epsilon_t \equiv A^{-1}B\eta_t$.
  - Zapisać model VAR w postaci VMA(∞)

$$y_t = \mu + \epsilon_t + \sum_{i=1}^{\infty} \Theta_i \epsilon_{t-k}$$

  - Ponieważ $\epsilon_t = A^{-1}B\eta_t$, IRF otrzymujemy jako

$$\Psi_i = \Theta_i A^{-1}B$$

Funkcja skumulowanej reakcji na impuls

- Niekiedy pracujemy na zmiennych typu ΔPKB_t, a interesuje nas reakcja PKB_t, nie pierwszej różnicy.
- Możemy otrzymać skumulowany IRF:

$$AIRF_{k(ij)} = \frac{\partial z_{i,t+k}}{\partial \eta_{jt}} = \sum_{m=0}^{k} IRF_{k(ij)}$$

FEVD

- Wariancja błędu losowego prognozy dla y_t dla k okresów do przodu to

$$\Sigma_k = Var\left(y_{T+k} - y^f_{T+k}\right) = \sum_{m=0}^{k-1} \Psi_m \Psi'_m$$

- Dla y_{it} na k okresów do przodu to

$$\Sigma_{k(ii)} = \sum_{j=1}^{N} \left(\Psi^2_{0(ij)} + \Psi^2_{1(ij)} + \ldots + \Psi^2_{k=1(ij)}\right)$$

- FEVD *(forecast error variance decomposition)* określa jaka część $\Sigma_{k(ii)}$ wynika z poszczególnych szoków strukturalnych

$$FEVD_{k(ij)} = \frac{\Psi^2_{0(ij)} + \Psi^2_{1(ij)} + \ldots + \Psi^2_{k=1(ij)}}{\Sigma_{k(ii)}}$$

FEVD

- Dla $k \to \infty$ FEVD można interpretować jako dekompozycję wariancji na zmiennych endogenicznych modelu VAR.
    - "21% zmienności inflacji można tłumaczyć szokami podażowymi, 12 szokami polityki pieniężnej..."

Strukturalizacja

- Jak otrzymać parametry modelu SVAR z parametrów modelu VAR?
- Model VAR

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \epsilon_t$$
$$\epsilon_t \sim N\left(0, \Sigma\right)$$

łącznie $K_{VAR} = PN^2 + N + \frac{N(N+1)}{2}$ parametrów

- Model SVAR

$$Ay_t = C_0 + C_1 y_{t-1} + C_2 y_{t-2} + \ldots + C_p y_{t-p} + B\eta_t$$
$$\eta_t \sim N\left(0, I\right)$$

łącznie $K_{SVAR} = N^2 + PN^2 + N + N^2$ parametrów

Strukturalizacja

- Potrzeba K dodatkowych restrykcji

$$
\begin{aligned}
K &= K_{SVAR} - K_{VAR} \\
&= N^2 + \frac{N(N-1)}{2}
\end{aligned}
$$

- Decyzja zależy od badacza, różne możliwe restrykcje:
    - strukturalizacja krótkookresowa (np. szok stopy procentowej wpływa na pewne zmienne z opóźnieniem)
    - strukturalizacja długookresowa (np. w długim okresie neutralność pieniądza)
    - *sign restrictions* (np. dodatni szok stopy procentowej zwiększa koszt kredytu dla firm)
    - ...
- Przyjmijmy od teraz, że $A = I$.

Strukturalizacja

- Rozważmy prosty przykład: niech y_t to PKB a r_t to stopa procentowa. Postać strukturalna:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \overbrace{\eta_t^y}^{\text{szok popytowy}} \\ \underbrace{\eta_t^r}_{\text{szok polityki pieniężnej}} \end{bmatrix}$$

  Postać zredukowana:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

- Pytanie badawcze: *jaki jest wpływ szoku polityki pieniężnej na PKB?*
- Nie możemy po prostu znaleźć funkcji reakcji na impuls ϵ_{2t} bo

$$\epsilon_{2t} = b_{21}\eta_t^y + b_{22}\eta_t^r$$

Strukturalizacja

- Czym jest ϵ_{2t}?

$$\epsilon_{2t} = b_{21}\eta_t^y + b_{22}\eta_t^r$$

  więc może to być zarówno szok polityki pieniężnej (przez b_{22}), jak i odpowiedź polityki pieniężnej na szok popytowy (przez b_{21}).

Strukturalizacja

- Problem identyfikacji polega na znalezieniu takiej macierzy B, że

$$\epsilon_t = B\eta_t$$

- Możemy w tym celu wykorzystać zależność, że

$$\Sigma = E\left[\epsilon_t \epsilon_t'\right] = BB'$$

- Problem: jest nieskończenie wiele macierzy $\tilde{B} = BZ$, gdzie $ZZ' = I$, które są rozwiązaniem.
- Intuicyjnie: Σ jest symetryczną macierzą, przez co mamy mniej równań, niż niewiadomych.
- Skąd wziąć brakujące równania?

Strukturalizacja

- W naszym prostym przykładzie:

$$
\begin{aligned}
\sigma_1^2 &= b_{11}^2 + b_{12}^2 \\
\sigma_{12} &= b_{11}b_{12} + b_{12}b_{22} \\
\sigma_{21} &= b_{11}b_{12} + b_{12}b_{22} \\
\sigma_2^2 &= b_{21}^2 + b_{22}^2
\end{aligned}
$$

czyli 3 równania i 4 niewiadome.

Strukturalizacja

- Przyjmimy, że polityka pieniężna działa na PKB z opóźnieniem (dlaczego?).
- To oznacza, że

$$\epsilon_{1t} = b_{11}\eta_t^y + \underbrace{b_{12}}_{=0}\eta_t^r,$$

czyli $b_{12} = 0$. Mamy układ równań

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ - & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ 0 & b_{22} \end{bmatrix}$$

Strukturalizacja

- Rozwiązanie:

$$b_{11} = \sigma_1$$

$$b_{21} = \frac{\sigma_{12}}{\sigma_1}$$

$$b_{22} = \sqrt{\sigma_2^2 - \frac{\sigma_{12}^4}{\sigma_1^2}}$$

Strukturalizacja

- To pozwala nam wyznaczyć efekt szoku polityki pieniężnej $\eta_t^r = 1$ jako

$$y_t = 0$$
$$r_t = \sqrt{\sigma_2^2 - \frac{\sigma_{12}^4}{\sigma_1^2}}$$

w bieżącym okresie. Resztę funkcji reakcji na impuls otrzymujemy w standardowy sposób.

Strukturalizacja

- Powyższy przykład to *rekursywna* strukturalizacja krótkookresowa (Sims, 1980).

$$B = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ * & * & \cdots & * \end{bmatrix}$$

  Idea bazuje na tym, że pewne zmienne działają na inne z opóźnieniem.
- Parametry można wyznaczyć rozwiązując układ równań

$$\Sigma = BB'$$

- Dekompozycja Cholesky'ego.