

Narzędzia modelowania przestrzennego.

dr Dorota Ciołek
Katedra Ekonometrii
Wydział Zarządzania UG

<http://wzr.pl/~dciolek>

Testowanie autokorelacji przestrzennej

autokorelacja globalna – zależność przestrzenne w obrębie całego badanego obszaru;

autokorelacja lokalna – zależność przestrzenna danej zmiennej w konkretnej lokalizacji z wartościami zmiennej w obszarach sąsiednich.

Sposoby testowania autokorelacji globalnej:

- Statystyka Gamma;
- Statystyki *joint-count* – zliczające połączenia;
- **Statystyka Morana I** ;
- Statystyka Geary'ego C .

Sposoby testowania autokorelacji lokalnej:

- Lokalna statystyka Morana I_i ;
- Lokalne statystyki autokorelacji przestrzennej Getisa i Orda.

Statystyka Morana I

przestrzenna autokorelacja globalna wg schematu opisanego konkretną macierzą wag W (zakładając, że macierz W jest standaryzowana wierszami - $\sum_i \sum_j w_{ij} = n$):

$$I_W = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\mathbf{z}^T \mathbf{W} \mathbf{z}}{\mathbf{z}^T \mathbf{z}}$$

gdzie: $\mathbf{z}$ – wektor kolumnowy o elementach:

Statystyka Morana I - cd

- rozkład statystyki jest asymptotycznie normalny.
- stwierdza, czy sąsiadujące ze sobą obszary są bardziej do siebie podobne niż wynikałoby z losowego charakteru zjawiska – czyli **czy występuje autokorelacja przestrzenna** badanej zmiennej.
- pokazuje **trafność doboru macierzy W** . Jeżeli macierz W opisuje stan rzeczywisty, czyli duża waga odpowiada dużej korelacji, to wartość statystyki Morana jest **wysoka**.

Macierze wag W

Zadaniem tzw. macierzy wag jest odzwierciedlenie relacji przestrzennych pomiędzy obiektami (krajami, województwami, powiatami, ale też miastami, dzielnicami, przedsiębiorstwami).

Budowa macierzy wag:

Etap 1: Macierz zero-jedynkowa wynikająca z geograficznych relacji pomiędzy sąsiadami

Etap 2: Standaryzacja macierzy – suma elementów w wierszu jest równa 1.

Macierze wag W - cd

Rodzaje macierzy wag:

- 1) Macierz oparta na kontyngencji stopnia pierwszego – kryterium wspólnej granicy.
- 2) Macierz k najbliższych sąsiadów.
- 3) Macierz sąsiadów w zasięgu d kilometrów (np. $d=50$)
- 4) Macierz wag przestrzennych rozumianych jako odwrotności odległości pomiędzy regionami (pomiędzy środkami regionów lub pomiędzy stolicami regionów).

Macierze wag W - cd

Macierz wag przestrzennych W (*spatial weight matrix*) definiująca strukturę przestrzennego sąsiedztwa, czyli przestrzenne powiązania i bliskość obserwacji.

Najczęściej wykorzystywana: **macierz najbliższego sąsiedztwa**, czyli macierz sąsiedztwa pierwszego rzędu, w której:

- $w_{ij} = 1$, gdy województwo i jest sąsiadem województwa j , tj. mają wspólną granicę;
- $w_{ij} = 0$, gdy województwo i nie jest sąsiadem województwa j , tj. nie mają wspólnej granicy;
- $w_{ii} = 0$, elementy diagonalne macierzy, czyli zakładamy, że województwo nie jest swoim własnym sąsiadem.

Modele regresji przestrzennej

1) **Modele Autoregresji Przestrzennej (SAR)– Modele Opóźnienia Przestrzennego (SLM)** (Arbia, 2006, s.110; Suchecki, 2010, s.248):

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

gdzie: $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$

$\mathbf{W}$ – macierz wag opisująca relacje przestrzenne pomiędzy jednostkami.
Model w postaci zredukowanej:

$$\mathbf{y} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\xi}$$

gdzie macierz wariancji i kowariancji składnika losowego:

$$E(\boldsymbol{\xi} \boldsymbol{\xi}^T) = \sigma^2 \left((\mathbf{I} - \rho \mathbf{W})^{-1} \left((\mathbf{I} - \rho \mathbf{W})^{-1} \right)^T \right)$$

Modele regresji przestrzennej - cd

2) Modele Autokorelacji Przestrzennej Składowika Losowego (SEM) (Arbia, 2006, s.112; Suček, 2010, s.250):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$$
$$\boldsymbol{\xi} = \lambda \mathbf{W}\boldsymbol{\xi} + \boldsymbol{\varepsilon}$$

W postaci zredukowanej:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda \mathbf{W})^{-1} \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$$

Modele regresji przestrzennej - cd

3) Modele z przestrzenną filtracją zmiennych objaśniających (modele regresji krzyżowych) (SCM)

(Arbia, 2006, s.115; Suhecki, 2010, s.251):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{WX}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$$

Zakładamy, że wartości zmiennej objaśnianej zależą od wartości wybranych zmiennych objaśniających w danym regionie oraz od ważonych wartości tych samych zmiennych w sąsiednich regionach.