空間統計勉強会

第 1 講 導入および空間的自己相関の検定

加藤真大

Tokyo, Japan

June 12, 2016

Contents

隣接性に着目した空間重み行列

例として隣り合った地区の影響を受けていることを次の方程式で示す。

$$\begin{cases} y_1 = f(X_1) + \rho y_2 \\ y_2 = f(X_2) + \rho(y_1 + y_3 + y_4 + y_5) \\ y_3 = f(X_3) + \rho(y_2 + y_4) \\ y_4 = f(X_4) + \rho(y_2 + y_3 + y_5) \\ y_5 = f(X_5) + \rho(y_2 + y_4) \end{cases} \tag{1}$$

隣接性に着目した空間重み行列

例として隣り合った地区の影響を受けていることを次の方程式で示す。

空間重み行列

$$c_{i,j} = \begin{cases} 1 & if\ i\ shares\ the\ same\ boundary\ with\ j \\ 0 & otherwise \end{cases} \tag{2}$$

$[C_{i,j}]$ を要素とする対称な行列 C を考えると、

$$\mathbf{C} = \begin{Bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{Bmatrix} .\mathbf{y} = \mathbf{f}(\mathbf{X}) + \rho\mathbf{C}\mathbf{y} \tag{3}$$

その他の空間重み行列

- k 近傍法
- 閾値なしの距離の逆数
- 閾値ありの距離の逆数

空間重み行列の基準化

- 計算の過程で $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1})$ という項の計算が必要になる場合がある。
- 逆行列の存在を保証するために基準化を行う。

大域的自己相関の検定

a グローバル・モラン（Moran の I 統計量）

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})} \tag{4}$$

大域的自己相関の検定

b グローバル・ギアリー（Geary の C 統計量）

$$I = \frac{n-1}{2S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij}(y_i - y_j)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

回帰モデルの残差の空間的自己相関の診断

- 誤差項に空間的自己相関が存在する場合、OLS 推定量が一致性をもたず、回帰係数の t 値が過大評価されるという問題が発生する。

誤差項の相関

-

空間的自己相関の有無に関する統計学的な診断

- 並べ替え検定。
- 漸近正規性を仮定した上で Z 検定。

局所的自己相関の検定

a ローカル・モラン

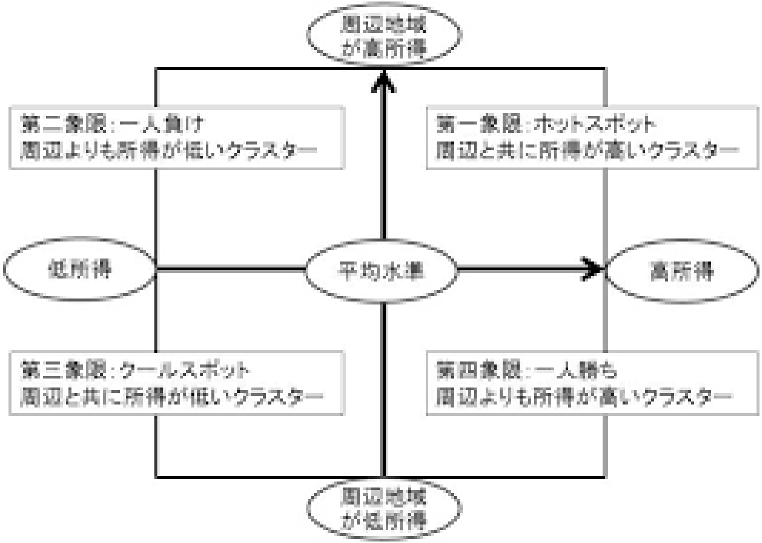
$$I_i = \frac{(y_i - y_j)}{m_2}\Sigma_{j=1}^{n}\omega_{ij}(y_i - \bar{y})I = \frac{\Sigma_i^n I_i}{S} \tag{6}$$

ランダム化仮説のもとでは,

$$E(I_i) = -\frac{\omega_i}{n-1} \tag{7}$$

$$\tag{8}$$

モラン散布図


周辺地域
が高所得
第二象限：一人負け
周辺よりも所得が低いクラスター
第一象限：ホットスポット
周辺と共に所得が高いクラスター
低所得
平均水準
高所得
第三象限：クールスポット
周辺と共に所得が低いクラスター
第四象限：一人勝ち
周辺よりも所得が高いクラスター
周辺地域
が低所得

局所的自己相関の検定

b G_i,$G*_i$ 統計量

$$G_i = \frac{\sum_{j \neq i} \omega_{ij} y_j}{\sum_{j \neq i} y_i} \tag{9}$$

$$G*_i = \frac{\sum_{j} \omega_{ij} y_j}{\sum_{j} y_i} \tag{10}$$

自身も統計量に含めるか否かで使い分ける。