

研究まとめ
「岡山市空き巣データの空間回帰モデリング」

大下 祐樹

2007年10月15日

1 研究動機

私は現在岡山の空き巣データと地域的指標の関係を解析することによって、地域分析を行おうとしている。現在までの解析手法として通常の回帰分析（OLS）、空間自己回帰モデル（SAR）、地理的加重回帰法（GWR）の3つのモデルを用いてきた。

本研究に用いるデータは地理データであり、観測値の空間的な従属性を考慮したモデルがSAR、GWRである。

それぞれ3つのモデルの推定パラメータ、モデル評価、誤差の独立性の結果を以降示す。

なお、地域的指標として老人化率、低層共同住宅割合を用いた。

2 解析結果

モデル評価には pss 、誤差の独立性については誤差の Moran'I 統計量、GWR の結果については要約等計量を示す。

$$pss = \sum_{i=1}^n (y_i - y_{(i)})^2 \quad (1)$$

$$\text{Moran'I} = \frac{1}{C} \frac{\sum_{i=1}^n \sum_{j=1}^n c_{ij}^* (y_i - \bar{y})(y_j - \bar{y})}{\sigma_y^2} \quad (2)$$

$$c_{ij} = \begin{cases} 1 & (\text{地区 } i \text{ と地区 } j \text{ が接するとき}) \\ 0 & (\text{上記以外}) \end{cases}$$

$$c^* = \frac{c_{ij}}{\sum_{j=1}^n c_{ij}}$$

$$C = \sum_{i=1}^n \sum_{j=1}^n c_{ij}^*$$

2.1 老人化率を説明変数

空き巣発生率（1000世帯あたり）と老人化率には以下のような負の相関関係が認められる。

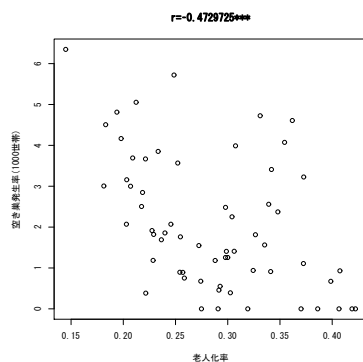


図 1: 岡山市：老人化率 vs. 空き巣発生率

2.1.1 老人化率を説明変数

$$\text{空き巣発生率}_i = \beta_0 + \beta_1 \cdot \text{老人化率}_i + \epsilon_i \quad (3)$$

単回帰の結果を以下に示す。

パラメータ	推定値	標準誤差	p 値
β_0	5.3236	0.8012	1.01e-08
β_1	-11.3888	2.7389	0.000104

残差二乗和	124.4839
pss	132.7853
残差の Moran'I	0.0841

2.1.2 SAR モデル (老人化率)

$$\text{空き巣発生率}_i = \beta_0 + \beta_1 \cdot \text{老人化率}_i + \rho \sum_{j=1}^n w_{ij} y_j + \epsilon_i \quad (4)$$

SAR モデルの結果を以下に示す。

パラメータ	推定値	標準誤差	p 値
β_0	4.637203	0.8336893	0.00000
β_1	-10.391553	2.8501405	0.00089
残差二乗和	121.4812		
pss	133.1101		
残差の Moran'I	0.004		
空間自己相関係数	0.182		

2.1.3 GWR モデル (老人化率)

$$\text{空き巣発生率}_i = \beta_{0i} + \beta_{1i} \cdot \text{老人化率}_i + \epsilon_i \quad (5)$$

GWR モデルの結果を以下に示す。

パラメータ	最小値	第 1 四分位	中央値	平均値	第 3 四分位	最大値
β_{0i}	4.377	5.021	5.405	5.382	5.577	6.855
β_{1i}	-17.260	-11.420	-11.000	-11.200	-10.320	-7.841
残差二乗和	112.2533					
pss	129.516					
残差の Moran'I	-0.0005					

2.2 低層共同住宅割合を説明変数

空き巣発生率（1000世帯あたり）と低層共同住宅割合には以下のような正の相関関係が認められる。

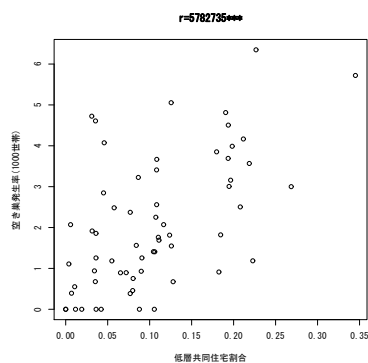


図 2: 岡山市：低層共同住宅割合 vs. 空き巣発生率

2.2.1 単回帰（低層共同住宅割合）

$$\text{空き巣発生率}_i = \beta_0 + \beta_1 \cdot \text{低層共同住宅割合}_i + \epsilon_i \quad (6)$$

単回帰の結果を以下に示す。

パラメータ	推定値	標準誤差	p 値
β_0	0.7975	0.2886	0.00759
β_1	12.2849	2.2375	8.56e-07
残差二乗和	106.7330		
pss	113.9339		
残差の Moran'I	-0.001		

2.2.2 SAR モデル（低層共同住宅割合）

$$\text{空き巣発生率}_i = \beta_0 + \beta_1 \cdot \text{低層共同住宅割合}_i + \rho \sum_{j=1}^n w_{ij} y_j + \epsilon_i \quad (7)$$

SAR モデルの結果を以下に示す。

パラメータ	推定値	標準誤差	p 値
β_0	0.7377761	0.2892056	0.01718
β_1	12.0519375	2.2423212	0.00000
残差二乗和	106.6240		
pss	116.9233		
残差の Moran'I	-0.0141		
空間自己相関係数	0.038		

2.2.3 GWR モデル（低層共同住宅割合）

$$\text{空き巣発生率}_i = \beta_{0i} + \beta_{1i} \cdot \text{低層共同住宅割合}_i + \epsilon_i \quad (8)$$

GWR モデルの結果を以下に示す。

パラメータ	最小値	第 1 四分位	中央値	平均値	第 3 四分位	最大値
β_{0i}	0.7781	0.8062	0.8137	0.8129	0.8189	0.8511
β_{1i}	12.00	12.15	12.19	12.19	12.23	12.42
残差二乗和	106.1911					
pss	113.9553					
残差の Moran'I	-0.0043					

	残差二乗和	pss	残差の Moran'I
OLS	124.4839	132.7835	0.0841
SAR	121.4812	133.1011	0.004
GWR	112.2533	129.516	-0.0005

表 1: 老人化率

	残差二乗和	pss	残差の Moran'I
OLS	106.7330	113.9339	-0.001
SAR	106.6240	116.9233	-0.0141
GWR	106.1911	113.9553	-0.0043

表 2: 低層共同住宅割合

3 まとめ

モデル評価の値をまとめは上記の表となる。低層共同住宅割合を説明変数にしたときに、GWR のガウスカーネルの最適バンド幅は求まらなかった。従って、低層共同住宅割合を説明変数にしたときは GWR は向いていないと言える。

それに比べ、最適バンド幅の求まった老人化率の場合では残差二乗和、pss 共に良化しており OLS に比べて被説明変数と説明変数の空間的変動をうまく考慮できたモデルになっていると言える。特に注目したいのが残差の Moran'I 統計量で、GWR では桁 1 つ小さくなっており、空間的に誤差が独立なモデルとなっている。ただし、OLS の場合でも充分小さい Moran'I 統計量である。