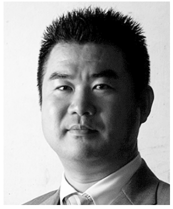
空間の統計学（2）：空間的自己相関

慶應義塾大学総合政策学部准教授
古谷 知之（Furutani Tomoyuki）
■兵庫県生まれ。2001年東京大学大学院工学系研究科博士課程修了。博士（工学）。東京大学大学院助手、慶應義塾大学環境情報学部専任講師を経て、07年4月より現職。専門分野：空間統計学、都市交通計画、観光政策。



1. はじめに

今回のテーマは、「空間的自己相関」です。自己相関（Autocorrelation）は、時系列解析において、よく知られる概念ですが、時系列自己相関のアナロジーから、空間データの系列相関を扱う際にも用いられます。空間的自己相関とは、おおまかにいえば、空間データの属性が、互いに近い地域・地点同士で似たような値を示す傾向があるか、それともランダムに分布しているかを示す指標です。点オブジェクトや面オブジェクトに対して、このような分析が適用されます。

空間的自己相関を計算するためには、「空間隣接行列（Spatial Neighbors Matrix、Neighborhood Matrix、Proximity Matrix）」と「空間重み付け行列（Spatial Weight Matrix）」を用います。空間隣接行列は、空間オブジェクト間の隣接関係を表します。空間重み付け行列は、空間オブジェクト間の隣接関係や距離をもとに、「近さ」を表現するための指標です。

空間的自己相関を示す指標には、グローバルな空間的自己相関を示す指標（Moran's Iなど）とローカルな空間的自己相関を示す指標（Geary's C、Local Moran's Iなど）があります。Rで空間的自己相関分析を行うために、今回は、spdepパッケージを利用します。

2. 空間隣接行列

空間オブジェクトが「隣り合っている」かどうかを表す指標として、空間隣接行列が用いられます。隣接関係（近接関係）の表現方法は、データの種類（メッシュデータ（グリッドデータ）、点データ、面データ）によって少しずつ異なります。以下では、まず、メッシュデータでの隣接関係を表現する方法を解説します。次に、点データや面データでの隣接関係を表現する際に用いられる方法として、ドロネー三角網、隣接地点（地区）数を指定する方法、一定半径以内に含まれる地点（地区）について隣接関係を定義する方法を、それぞれ説明します。

⑴　メッシュデータ

まず、メッシュデータの隣接関係の表現方法からみてみましょう。メッシュデータの場合、例えばチェスのアナロジーで、図１のように隣接性を決める方法があります。図１のうち、⒜のように上下左右に接するメッシュのみ隣接関係を定義する場合をルーク型隣接関係（Rook Type Neighborhood）、⒝のように角に接するメッシュのみ隣接関係を定義する場合をエッジ型隣接関係（Edge Type Neighborhood）、⒞のように上下左右と角が接するメッシュの隣接関係を定義する場合をクイーン型隣接関係（Queen Type Neighborhood）などと呼ぶことがあります。

次に、単純なメッシュデータの例として、図２のような3×3のグリッド状の９区画を取り上げ、ルーク型の隣接関係を空間隣接行列として表現してみましょう。このとき、当該メッシュに対して上下左右の辺が接しているメッシュについては「隣接している（＝1）」、そうでない場合は「隣接していない（＝0）」と定義することにすると、空間隣接行列Cは、次式のように表すことができます。

$$C=\begin{pmatrix} 0&1&0&1&0&0&0&0&0\\ 1&0&1&0&1&0&0&0&0\\ 0&1&0&0&0&1&0&0&0\\ 1&0&0&0&1&0&1&0&0\\ 0&1&0&1&0&1&0&1&0\\ 0&0&1&0&1&0&0&0&1\\ 0&0&0&1&0&0&0&1&0\\ 0&0&0&0&1&0&1&0&1\\ 0&0&0&0&0&1&0&1&0 \end{pmatrix} \qquad (1)$$

図１　隣接関係の定義例

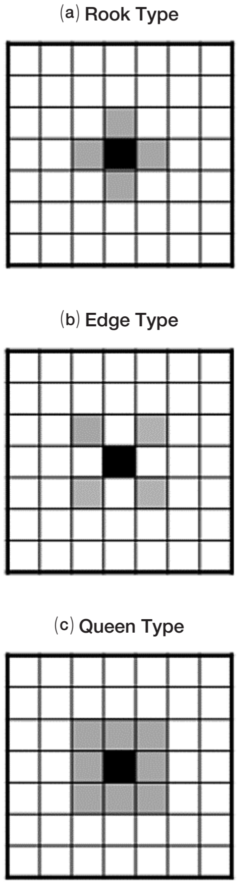
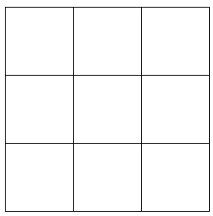


図２　３×３グリッドデータ



⑵　ドロネー三角網

点データの場合、「最も近い地点（地区）」同士を結んで隣接関係を定義することができます。面データの場合でも、ポリゴンの代表

点を定義することにより、点データと同じように隣接関係を定義できるといえます。

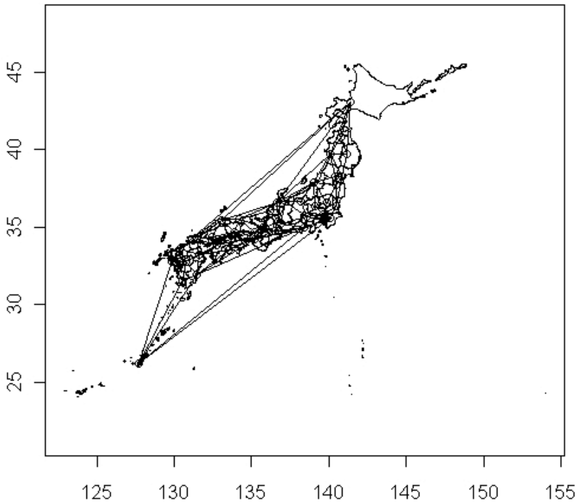
地点同士をつなぎ、隣接関係を示す方法の一つに、ドロネー三角網（Delaunay Triangulation）という方法があります。大雑把にいえば、当該点から近い点同士を結んで三角形を形成する方法です。

Rでは、緯度経度データを用いて、**tri2nb( )**関数によりドロネー三角網による隣接関係を定義できます。ここでは、日本の都道府県庁所在地の緯度経度データを用いて、ドロネー三角網図を作成してみましょう。

```
library(spdep)
pref.pnt <- read.shape("pref_gov.shp")
jpn.pref <- Map2poly(jpn.pref)
jpn.pref <- read.shape("jpn_pref.shp")
pref.pnt <- Map2points(pref.pnt)
plot(jpn.pref)
points(pref.pnt)
pref_gov <- read.table(
"pref_gov.txt",",",header=T,
row.names=2)
coords <- matrix(0,nrow(pref_gov),2)
coords[,1] <- pref_gov$X
coords[,2] <- pref_gov$Y
pref.tri.nb <- tri2nb(coords,
row.names=rownames(pref_gov))
plot(pref.tri.nb, coords, add=T)
```

図３のように、ドロネー三角網を使った場合は、都道府県境界が隣接しているかどうかにかかわらず、隣接関係が定義されます。また、サンプリングポイントが多いエリアは、より密にネットワークが形成されることがわかります。

図３　ドロネー三角網図



⑶　近隣 k 地点に隣接関係を定義する方法

当該オブジェクトから近接する地点数をあらかじめ決めておき、隣接関係を定義する方法があります。この方法は、各地点間の距離をもとに、最も近い k 地点のみを抽出し、それらを近接した地点と定義する方法です。

Rでは、座標行列を用いて**knearneigh( )**関数により近隣 k 地点の隣接関係を定義することができます。この方法を用いた場合、ドロネー三角網図と比較して、より近い地点同士の近接関係を定義することができるといえます（図４）。**diffnb( )**関数を使って、$k=3$とした場合との違いを示すこともできます。

```
pref.knn <- knearneigh(coords, k=4)
pref.knn.nb <- knn2nb(pref.knn)
plot(jpn.pref)
plot(pref.knn.nb, coords, add=TRUE)
```

図４　近接 k 地点の隣接関係（k＝４の場合）

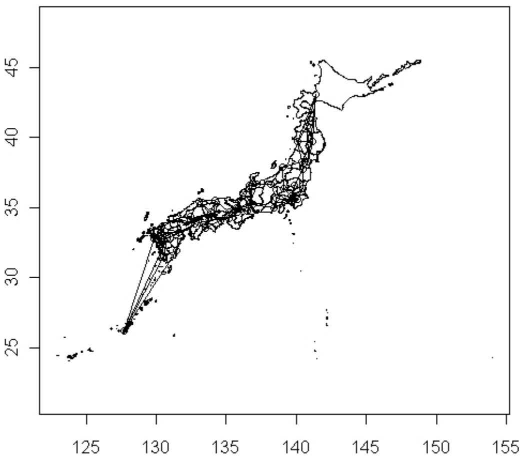
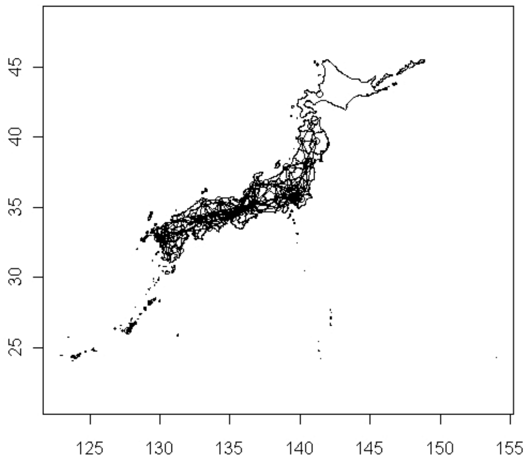


図５　ユークリッド距離による隣接関係（r＝２の場合）



⑷　ユークリッド距離により隣接関係を定義する方法

ある一定のユークリッド距離 r をあらかじめ決めておき、半径 r の円の内部に含まれる地点についてのみ隣接関係を定義する方法です。また、その応用として、上限値 r_1 を半径とする円のうち、下限値 r_2 を半径とする円を除いた領域に含まれる地点についてのみ隣接関係を定義する方法もあります。

Rでは**dnearneigh( )**関数により、ユークリッド距離による隣接関係を定義することができます。半径 r 以内に近接地点が一つも含まれない地点は、隣接関係が定義されないことになります。例えば、$r=2$ として計算した場合、北海道と沖縄には隣接関係が定義されません（図５）。

```
pref.r.nb <- dnearneigh(coords, 0, 2,
row.names=rownames(pref_gov))
plot(jpn.pref)
plot(pref.r.nb, coords, add=TRUE)
```

⑸　ポリゴンデータの隣接関係

ポリゴンデータについては、ポリゴンの境界同士が隣接しているかどうかで隣接関係を定義する方法があります。Rでは、**poly2nb( )**関数により、ポリゴンデータの隣接関係を定義することができます。

3. 空間重み付け行列

隣接行列が定義されれば、隣接関係を重み付け化して地点（地区）間の近接性を表現できます。これを、空間重み付け行列（空間重み行列、Spatial Weight Matrix）といいます。空間重み付け行列には、⑴隣接行列をそのまま用いる場合（W_B）、⑵隣接行列の行和で標準化する場合（W_W）、⑶行列の全要素の和で標準化する場合（W_C）、⑷距離行列を用いる場合（W_S）、などがあります。

ふたたび、図２のようなグリッド上の９区画を対象に空間重み付け行列を作ることを考えてみましょう。要素 c_{ij} $(i, j=1, \cdots, N)$ からなる空間隣接行列 C が与えられたとき、隣接

行列の行和で標準化するW_Wの要素W_{ij}は次式のように表されます。

$$w_{ij}=\frac{c_{ij}}{\sum_{i=1}^{N}c_{ij}} \tag{2}$$

このとき、空間重み付け行列W_Wは、次式のようになります。

$$W_W=\begin{pmatrix} 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \end{pmatrix} \tag{3}$$

これをRで計算すると、次のようになります。

```
N=9
neigh.rook <- matrix(ncol=N,nrow=N,c(
0,1,0,1,0,0,0,0,0,
1,0,1,0,1,0,0,0,0,
0,1,0,0,0,1,0,0,0,
1,0,0,0,1,0,1,0,0,
0,1,0,1,0,1,0,1,0,
0,0,1,0,1,0,0,0,1,
0,0,0,1,0,0,0,1,0,
0,0,0,0,1,0,1,0,1,
0,0,0,0,0,1,0,1,0))
neigh.rook.w <- neigh.rook
for(i in 1:N){neigh.rook.w[i,] <-
neigh.rook[i,] / sum(neigh.rook[i,])}
```

4. 空間的自己相関分析

空間的自己相関を表す指標として、グローバルな空間的自己相関を示すMoran's Iと局地的な空間的自己相関を示すGeary's Cが知られています。Moran's Iは地区（地点）毎に計算することができ、局地的な空間的自己相関を示すLocal Moran's Iとして知られています。

⑴ Moran's I

N地点（地区）からなる対象地域において、地点iの属性をx_iとします。対象地域全体における属性の平均値を$\bar{x}$、地点ij間の空間重み付け行列の要素をw_{ij}とすると、Moran's Iは次式のように表すことができます。

$$I=\frac{N\sum_{i=1}^{N}\sum_{j=1}^{N}w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{N}\sum_{j=1}^{N}w_{ij}\sum_{i=1}^{N}(x_i-\bar{x})^2} \tag{4}$$

この式に示されているように、Moran's Iは相関係数に空間重み付け行列の要素を考慮した指標であるといえます。Moran's Iは、−1から1の間の数値をとります。1に近い値のとき、互いに近い空間オブジェクトの属性が類似するような、空間的自己相関が強いことを意味します。また、−1に近い値のときは、空間オブジェクトの属性がばらついているといえます。

Rでは、**moran.test**( )関数を使ってMoran's Iを計算することができます。またこのとき、**nb2listw**( )関数により空間重み付け行列を与えることができます。

```
moran.test(pref.pnt$diabetes,
nb2listw (pref.tri.nb,style="W"))
```

⑵　Geary's C

Geary's Cは次式のように表され、０から２の間の値をとります。Moran's Iとは異なり、０に近い値をとる場合は正の空間的自己相関を示し、２に近い値をとる場合は負の空間的自己相関を示します。また、１のときは、空間的自己相関がないといいます。

$$C = \frac{(N-1)\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(x_i - x_j)^2}{2W\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (5)$$

Rでは、**geary.test( )**関数を使ってGeary's Cを計算することができます。

```
geary.test(pref.pnt$diabetes,
nb2listw(pref.tri.nb, style="W"))
```

⑶　Local Moran's I

最後に、局地的な空間的自己相関を示すLocal Moran's Iを計算します。Local Moran's Iは、次式のように表されます。

$$I_i = \frac{(n-1)(x_i - \bar{x})}{\sum_{j}^{n}(x_j - \bar{x})^2} \cdot \sum_{j=1, j \neq 1}^{n} w_{ij}(x_j - \bar{x}) \qquad (6)$$

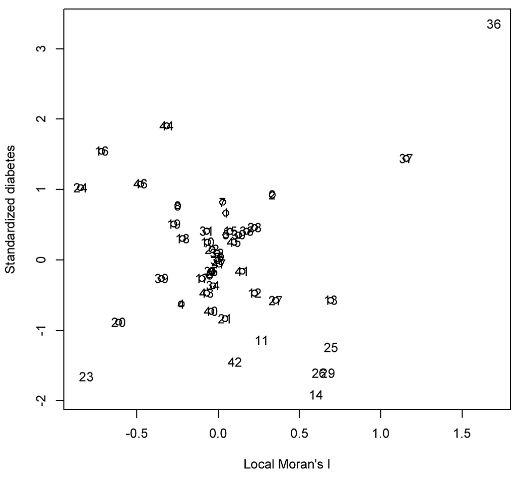
Rでは、**localmoran( )**関数を使ってLocal Moran's Iを計算することができます。

```
LMI1 <- localmoran(pref.pnt$diabetes,
nb2listw(pref.tri.nb, style="W"))
```

```
pref.lm <- data.frame(cbind(LMI1[,1],(pref.
pnt$diabetes-
mean(pref.pnt$diabetes))/sd(pref.pnt$diabetes)),
row.names=pref.pnt$KENCODE)
colnames(pref.lm) <- c("li","standardized
diabetes")
pref.lm
plot(pref.lm,xlab="Local Moran's I",
ylab="Standardized diabetes")
text(pref.lm,rownames(pref.lm))
```

Local Moran's Iと標準化された属性値をプロットしたとき、第一象限にプロットされる地区（地点）は、属性値が他の地区と比較して相対的に大きく、かつ類似する値を持つ（空間的に自己相関する）地区（地点）が周囲にあることを意味します（図６）。

図６　Local Moran's Iと標準化された属性値のプロット



＊参考URL
総務省統計局：社会生活統計指標－都道府県の指標2009（http://www.stat.go.jp/data/ssds/5.htm）.