2011年1月28日

第4章　空間解析
7.　空間的自己相関

井上 亮
rinoue@plan.civil.tohoku.ac.jp

ここで学ぶこと

- 空間的自己相関
  - 空間的自己相関とは
  - 空間的自己相関の統計量
    - Join 統計量
    - Moran’s I 統計量
    - Geary’s C 統計量
    - Getis-Ord G 統計量
  - 空間的自己相関のモデル化

について学びます．

空間的自己相関とは

- ある空間情報に
「（正の）空間的自己相関がある」とは
，
「近い点における属性の類似度が大きい」
ことを意味する．

空間的自己相関とは

気象条件

気温に
影響を与える要因

土地利用
土地被覆

観測点 A
30℃

観測点 B
26℃

観測点 C
18℃

近隣では要因が類似→気温に高い類似性

空間的自己相関とは

- ある空間情報に
「（正の）空間的自己相関がある」とは，
「近い点における属性の類似度が大きい」
ことを意味する．

一般に，空間データには，
（正の）空間的自己相関が存在することが多い．

地理学第一法則 (Tobler(1970))

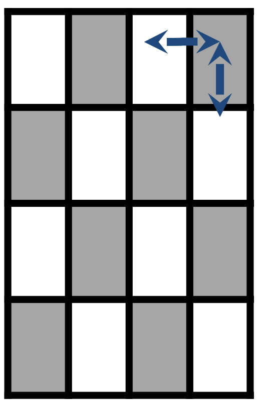
“Everything is related to everything else,
but near things are more related than distant things”

負の空間的自己相関？

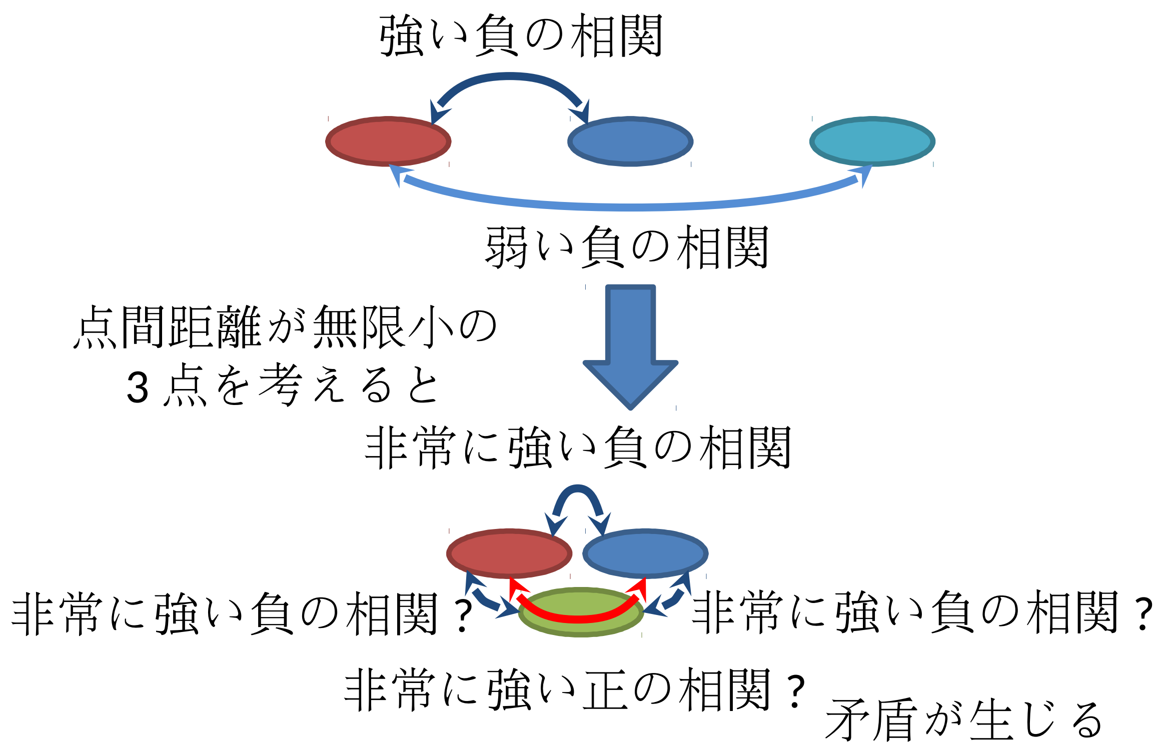
「負の空間的自己相関」：
空間的に近いほど属性は負の相関を持つ

離散的な空間

「隣り合う面が負の相関を持つ」
ことは可能



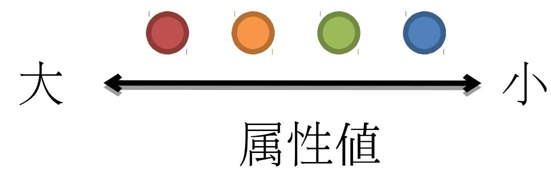
連続的な空間

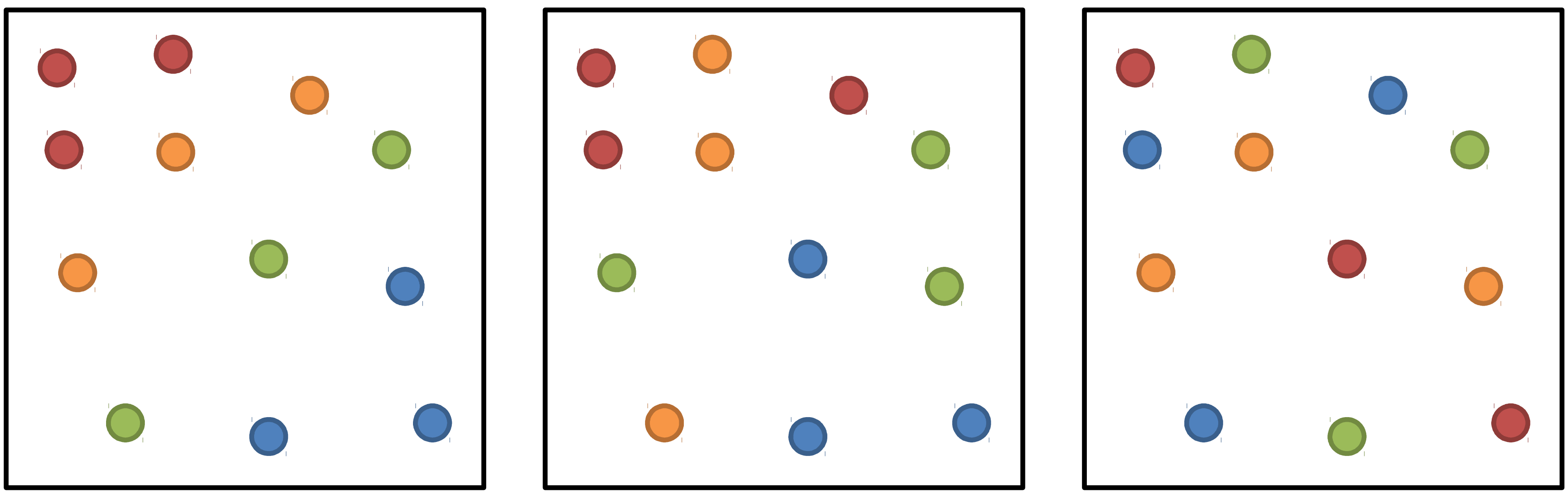


連続的な空間では
正の空間的自己相関のみを考える

空間的自己相関の定量化

- 相関の強弱





強い　弱い　無相関

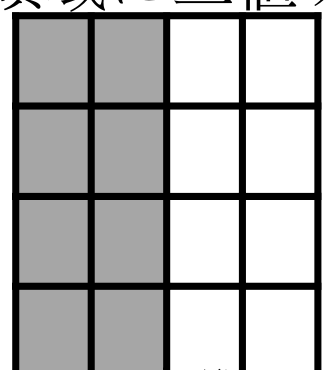
空間的自己相関を表す統計量

- Join 統計量
- Moran’s I 統計量
- Geary’s C 統計量
- Getis-Ord G 統計量

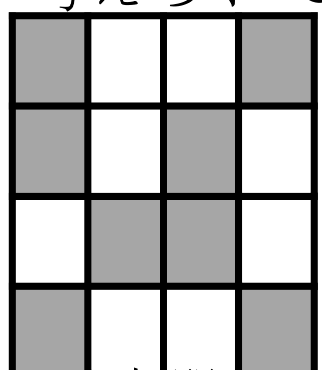
Join 統計量

- 離散的な領域（例えば，格子領域や面データ）に与えられた
  カテゴリデータ（例えば，二値データ）の
  空間的自己相関を表す統計量

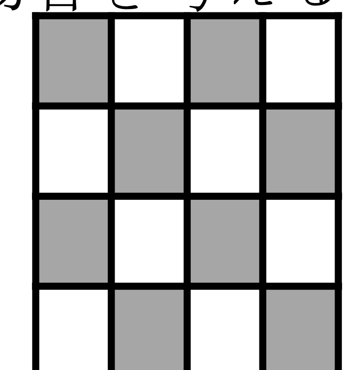
格子領域に二値データが与えられている場合を考える．



正の相関？



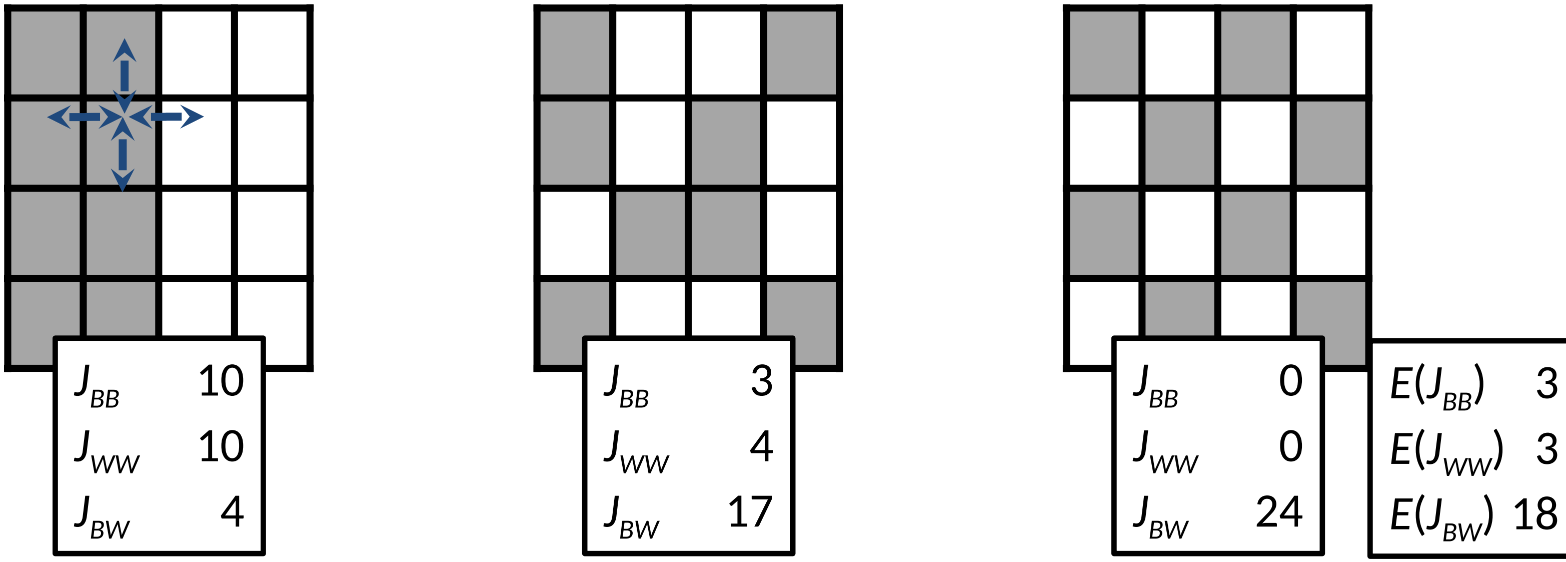
無相関？



負の相関？

Join 統計量

「格子領域の隣接関係を辺を共有する関係」と定義した場合，
黒黒の隣接ペア数 J_{BB}, 白白の隣接ペア数 J_{WW}, 黒白の隣接ペア数 J_{BW} を数える

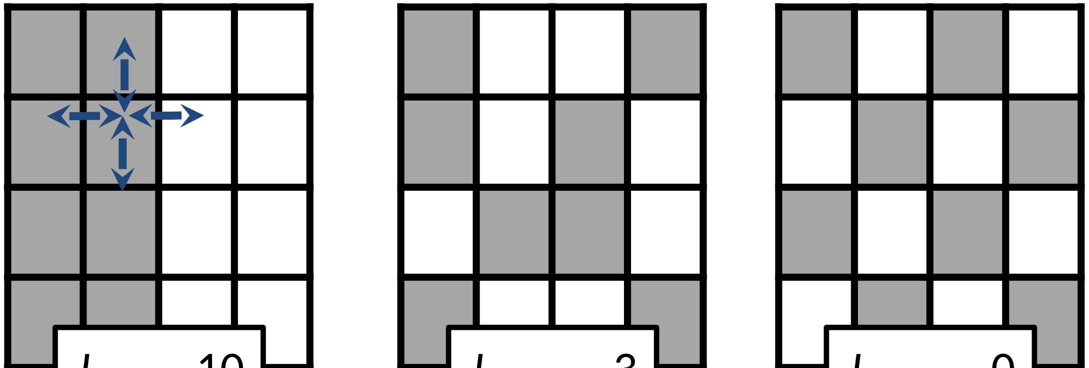


黒の発生確率を p_B，白の発生確率を p_W，総隣接ペア数を k とすると
ランダムに発生した (無相関の) 場合の隣接ペア数の期待値は

$$E(J_{BB}) = k\,p_B^2$$
$$E(J_{WW}) = k\,p_W^2$$
$$E(J_{BW}) = 2k\,p_B\,p_W$$

上図の場合， $p_B = p_W = 0.5$, $k=24$ なので， $E(J_{BB}) = E(J_{WW}) = 3$, $E(J_{BW}) = 18$.
中央は「ほぼ無相関」といえる．

Join 統計量



J_{BB}	10
J_{WW}	10
J_{BW}	4

J_{BB}	3
J_{WW}	4
J_{BW}	17

J_{BB}	0
J_{WW}	0
J_{BW}	24

$E(J_{BB})$	3
$E(J_{WW})$	3
$E(J_{BW})$	18

た， $m = \frac{1}{2}\sum_{i=1}^{n} k_i\left(k_i - 1\right)$ 　ただし k_i は i 番目領域の隣接数

すると， J_{BB}, J_{WW}, J_{BW} の分散 $s_{BB}^2, s_{WW}^2, s_{BW}^2$ の期待値は

$$E\left(s_{BB}^2\right) = kp_B^2 + 2mp_B^3 - \left(k + 2m\right)p_B^4$$

$$E\left(s_{WW}^2\right) = kp_W^2 + 2mp_W^3 - \left(k + 2m\right)p_W^4$$

$$E\left(s_{BW}^2\right) = 2\left(k + m\right)p_B p_W - 4\left(k + 2m\right)p_B^2 p_W^2$$

上図の場合， $p_B = p_W = 0.5$, k=24, m=52 なので， $E\left(s_{BB}^2\right) = E\left(s_{WW}^2\right) = 11$, $E\left(s_{BW}^2\right) = 6$

左図が「空間的自己相関を持ったデータ」であると言えるだろう．

領域数が多ければ，統計的検定を行うことが可能．

Join 統計量

なお，

z_i: 領域 *i* の二値データの値 (黒 : 1, 白 : 0),

w_{ij}: 隣接関係（領域 *i* と領域 *j* が隣接 : 1, 隣接していない : 0）

とし，ベクトル **z**=(z_1,…,z_n)', 隣接行列 **W**={w_{ij}} を定義とすると

$$J_{BB} = \frac{1}{2}\mathbf{z}'\mathbf{W}z$$

と表すことができる.

Join 統計量

限界

- 離散的な空間領域に基づく「隣接関係」を通してしか空間的自己相関の有無を判断できない．
  - 点データへの応用ができない
  - 隣接関係の定義に距離を反映させることができない

- 扱える属性はカテゴリデータのみ
  属性が連続量の場合，カテゴリデータに変換しないと適用できない．

離散的な空間領域に基づかず，点データにも適用が可能で，

より柔軟に空間的な近さ（隣接関係以外に距離など）を考慮でき，

連続量の属性が扱える統計量が必要！

Moran’s I 統計量

$$I=\frac{n}{\sum_i\sum_j w_{ij}}\frac{(\mathbf{z}-\bar{\mathbf{z}})'\mathbf{W}(\mathbf{z}-\bar{\mathbf{z}})}{(\mathbf{z}-\bar{\mathbf{z}})'(\mathbf{z}-\bar{\mathbf{z}})}$$

n: 領域 (点) の数

z : 属性値ベクトル（ $\mathbf{z}=(z_1,\ldots,z_n)'$, z_i : 領域 (点)i の属性値）

$\bar{\mathbf{z}}$: 属性値の平均値ベクトル（ $\bar{\mathbf{z}}=(\bar{z},\ldots,\bar{z})'$ $\bar{z}=\frac{1}{n}\sum_i z_i$ ）

$\mathbf{W}=\{w_{ij}\}$: 空間重み行列

ただし， w_{ij}: 領域 (点)i,j の近接関係を表す

隣接（隣接 : w_{ij} =1, 隣接していない : w_{ij} = 0 ）

あるいは

距離の逆数の定数乗（ $w_{ij}=d_{ij}^{-\alpha}$ ）ただし α ＞ 0

など

Moran’s I 統計量

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{(z-\bar{z})' \mathbf{W} (z-\bar{z})}{(z-\bar{z})'(z-\bar{z})}$$

Join 統計量と類似

$$= \frac{(z-\bar{z})' \mathbf{W} (z-\bar{z})}{\left\{\sum_i \sum_j w_{ij}\right\} \left\{\frac{1}{n}(z-\bar{z})'(z-\bar{z})\right\}}$$

空間重み行列の要素の和　　属性値の分散に相当

負の空間的自己相関 $-1 \leq I \leq 1$ 正の空間的自己相関

Moran’s I 統計量

$$I=\frac{n}{\sum_i\sum_j w_{ij}}\frac{(\mathbf{z}-\bar{\mathbf{z}})'\mathbf{W}(\mathbf{z}-\bar{\mathbf{z}})}{(\mathbf{z}-\bar{\mathbf{z}})'(\mathbf{z}-\bar{\mathbf{z}})}$$

属性値が正規分布に従っており，
空間的自己相関がないと仮定すると

期待値 $$E(I)=\frac{-1}{n-1}$$

分散 $$Var(I)=\frac{n^2\left(\sum_i\sum_j\left(w_{ij}+w_{ji}\right)\right)-n\left(\sum_i\left(w_{i\cdot}+w_{\cdot i}\right)\right)+3\left(\sum_i\sum_j w_{ij}\right)^2}{\left(\sum_i\sum_j w_{ij}\right)^2\left(n^2-1\right)}-E(I)^2$$

データ数が多い場合は，近似的に正規分布に従う
→空間的自己相関の有無の検定が可能

Geary's C 統計量

Moran's I 統計量と類似の統計量

$$C = \frac{n-1}{2\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} \left(z_i - z_j \right)^2}{\sum_i \left(z_i - \bar{z} \right)^2}$$

C>1　負の空間的自己相関
C=1　相関なし
C <1　正の空間的自己相関

Getis-Ord G 統計量

$$G_i = \frac{\sum_j w_{ij} z_j}{\sum_j z_j}$$

統計量の期待値

$$E(G_i) = \frac{\sum_j w_{ij}}{n - 1}$$

分散

$$Var(G_i) = \frac{\sum_j w_{ij}\left(n - \sum_j w_{ij}\right)s^2}{n^2(n-1)\bar{z}^2}$$

$$\bar{z} = \frac{\sum_i z_i}{n}, s^2 = \frac{\sum_i (z_i - \bar{z})^2}{n - 1}$$

空間的自己相関のモデル化

空間確率場 Z に対して，「本質的定常性」を仮定

本質的定常性　(intrinsic stationarity)：
空間確率場 Z 上の確率変数 Z($\mathbf{s}_i$)，　Z($\mathbf{s}_j$) の分散が，
空間的な相対位置だけで決まるとの仮定.

$$E\left[Z(\mathbf{s}_i)-Z(\mathbf{s}_j)\right]=0$$

$$Var\left(Z(\mathbf{s}_i)-Z(\mathbf{s}_j)\right)=2\gamma\left(\mathbf{s}_i-\mathbf{s}_j\right)$$

Z($\mathbf{s}_i$)：　点 $\mathbf{s}_i$ における確率変数

更に等方性（方向によって空間相関の影響が変わらない）を仮定すると

$$Var\left(Z(\mathbf{s}_i)-Z(\mathbf{s}_j)\right)=2\gamma\left(\left\|\mathbf{s}_i-\mathbf{s}_j\right\|\right)=\underline{2\gamma\left(d_{ij}\right)}$$

2 点間の距離だけの関数で分散を表す

バリオグラム

空間的自己相関のモデル化

空間確率場 Z に対して，「2 次定常性」を仮定．

2 次定常性　(Second order stationarity)：
空間確率場 Z 上の確率変数 $Z(\mathbf{s}_i)$，　$Z(\mathbf{s}_j)$ の共分散が，
空間的な相対位置だけで決まるとの仮定．

$$E\left(Z\left(\mathbf{s}_i\right)\right)=\mu$$

$$Cov\left(Z\left(\mathbf{s}_i\right),Z\left(\mathbf{s}_j\right)\right)=C\left(\mathbf{s}_i-\mathbf{s}_j\right)$$

$Z(\mathbf{s}_i)$：　点 $\mathbf{s}_i$ における確率変数

更に等方性（方向によって空間相関の影響が変わらない）を仮定すると

$$Cov\left(Z\left(\mathbf{s}_i\right),Z\left(\mathbf{s}_j\right)\right)=C\left(\left\|\mathbf{s}_i-\mathbf{s}_j\right\|\right)=\underline{C\left(d_{ij}\right)}$$

2 点間の距離だけの関数で共分散を表す

コバリオグラム（共分散関数）

バリオグラムとコバリオグラム

二次定常性の仮定の下では

$$2\gamma(\mathbf{s}_i - \mathbf{s}_j) = Var\{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\}$$

$$= Var(Z(\mathbf{s}_i)) + Var(Z(\mathbf{s}_j)) - 2Cov\{Z(\mathbf{s}_i), Z(\mathbf{s}_j)\}$$

$$= 2\{C(0) - C(\mathbf{s}_i - \mathbf{s}_j)\}$$

バリオグラム

・距離０では不連続（距離０では０，少しでも離れると正の値）

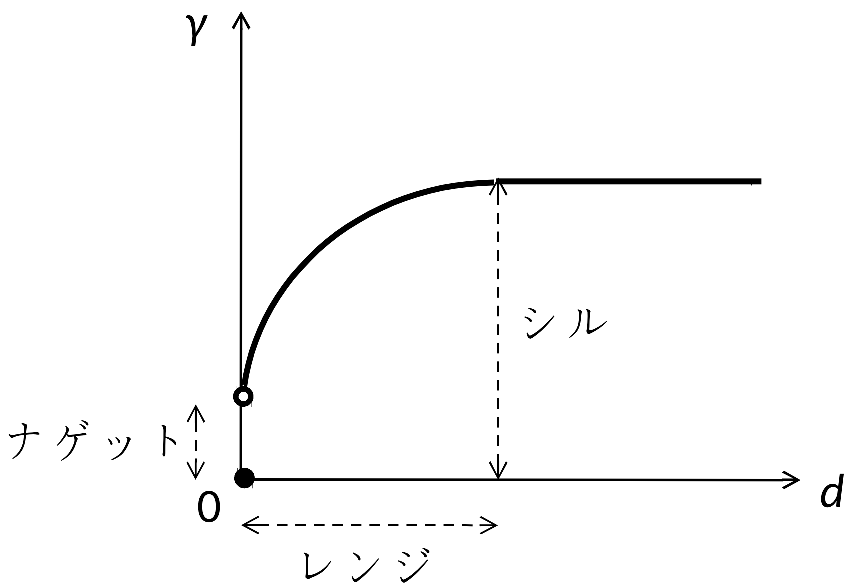
・距離が離れると分散は増加

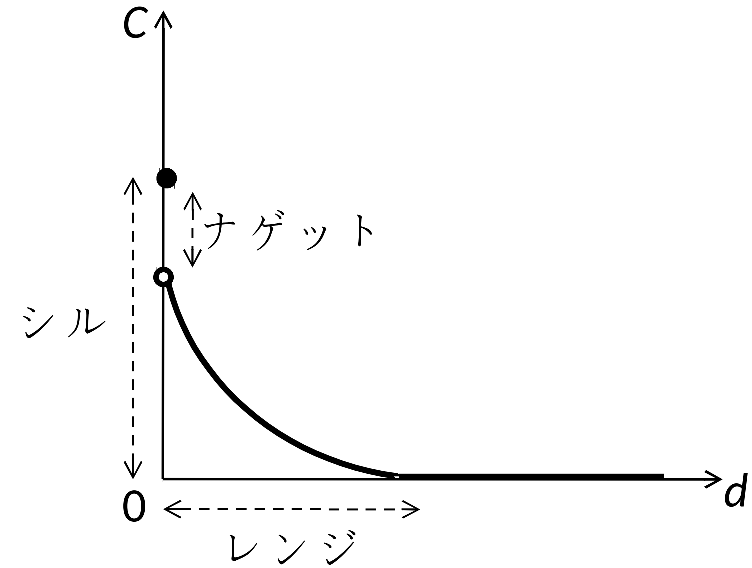
コバリオグラム（共分散関数）

・距離０では不連続（分散と共分散の値は連続ではない）

・距離減衰（離れるにつれて共分散 小）

バリオグラムとコバリオグラム





任意の地点間の空間的自己相関の影響の大きさを
距離の関数としてモデル化

参考文献

Cliff, A. D. & Ord, J. K.: *Spatial Processes –Models & Applications*, Pion, 1981.

Getis, A.: Spatial autocorrelation, Fischer, M. M.& Getis A. Eds.: *Handbook of Applied Spatial Analysis –Software Tools, Methods and Applications*, Springer, 2010.

Haining, R.: Spatial Data Analysis in the Social and Environmental Sciences, Cambridge University Press, 1990.

間瀬 茂・武田 純：空間データモデリング，共立出版, pp.135-166, 2001.

Cressie, N. A. C.： *Statistics for Spatial Data*, John Wiley & Sons, 1993.