第4章 空間解析 10 空間分析におけるスケール

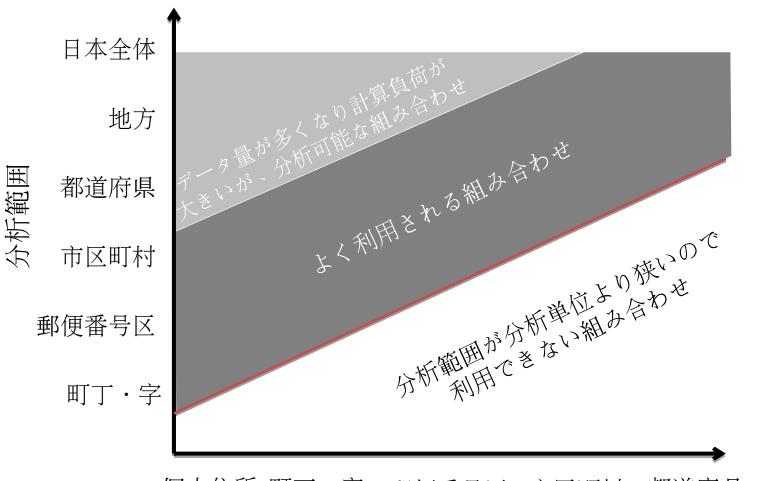
中谷友樹 nakaya@lt.ritsumei.ac.jp

ここで学ぶこと

■空間スケールは空間分析においてどのような意味をもつのだろうか?

- ◆可変地区単位問題
 - ◆分析単位の問題
- ◆ローカルとグローバル問題
 - ◆分析範囲の問題

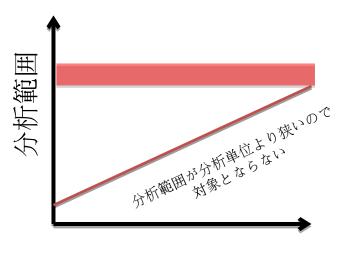
2つの空間スケール 空間単位と分析範囲



個人住所 町丁·字 郵便番号区 市区町村 都道府県 分析単位

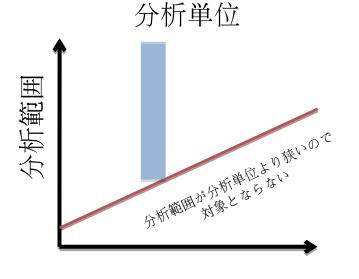
地理情報科学教育用スライド © 中谷友樹

2つの空間スケール問題



分析範囲は同じだが、分析単位が異なると 結果が変わりうる問題 (単位の定義の違い)

→ 可変空間単位問題

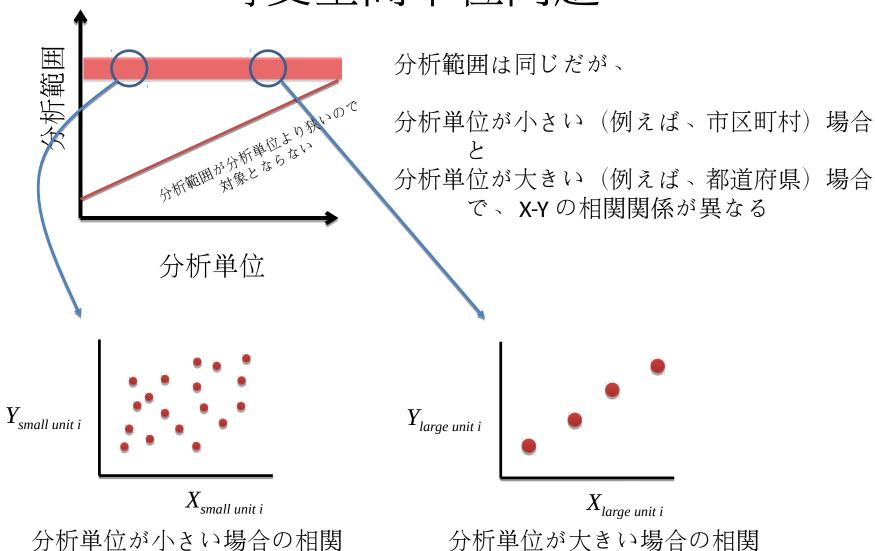


分析単位は同じだが、分析範囲が異なる ことで生じる問題(全体と部分の違い)

→ ローカル・グローバル問題

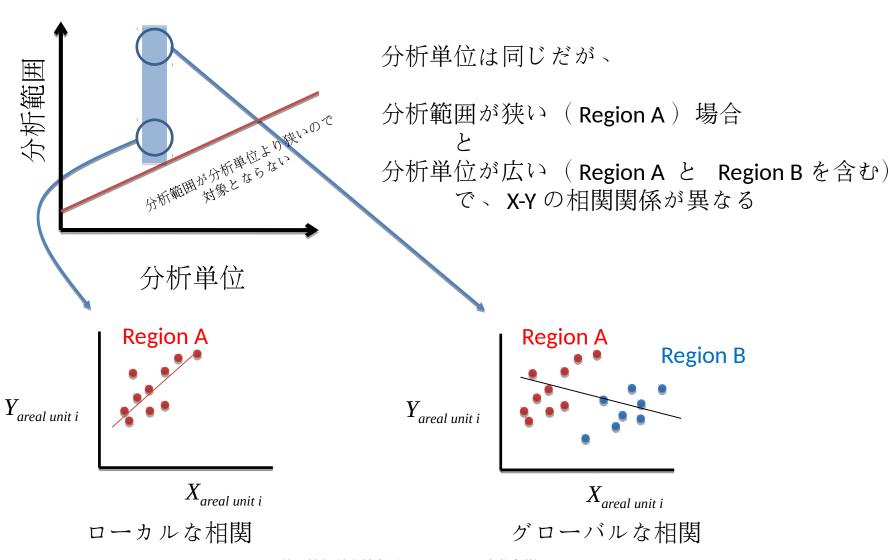
分析単位

2 変量相関の仮想事例 可変空間単位問題



地理情報科学教育用スライド © 中谷友樹

2変量相関の仮想事例 ローカルとグローバル問題



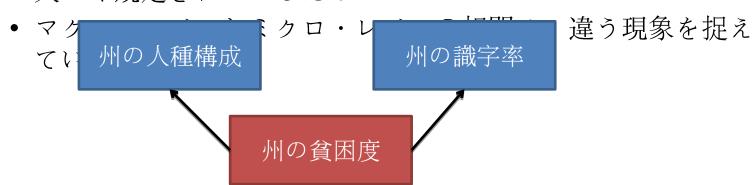
可変空間単位問題 Modifiable areal unit problem (MAUP)

可変空間単位

- ■人口密度、収穫率、投票率など、率指標は適当な地理的な単位で集計して得られる
- ■とくに率指標は、あるサンプル数が多くなる単位でないと統計的に不安定であるので、集計によるメリットがある
- ■しかし、地理的な単位は、いくらでも「変更可能 modfiable」な性質を持つ(Yule and Kendall, 1950)
- ■そのため、異なる地理的単位の利用が、異なる 分析結果をもたらすなら、分析結果の妥当性あ るいは一般化は保証されない。

生態学的誤謬 ecological fallacy

- マクロ・レベル(集計単位による)相関をミクロ・レベル(非集計単位による)相関と解釈する誤り
 - 社会学者 Robinson(1950) による問題提起
 - 個人レベルでは、人種と識字率に相関はない
 - 州レベルでは、非白人の割合が高い地域ほど、識字率は低い
 - 州という集計レベルでは、人種構成と識字率が州の貧困度に 大きく規定されているため。



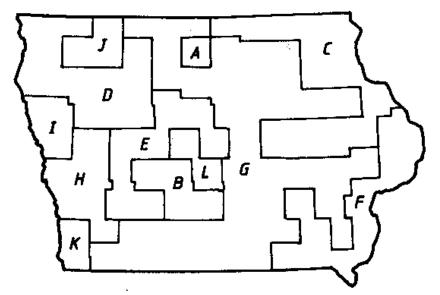
スケール問題とゾーニング問題

ゾーニング問題 集計方法間で単位間の関係が 一致するとみる推論の誤謬 推論の誤謬 非集計単位の関係を集計単位間 みる推論の誤謬 集計単位の関係を非集計 スケール問題 開霧 M 约 生態学[単位間の関係と の関係と 集計方法1 集計方法2

Openshaw の実験

- lowa 州 99 郡における、高齢者比率 (X)と共和党支持者への 投票率 (Y)の相関係数: 0.34
- 様々な方法で、より大きな地理的単位に 99 郡のデータを集計してみると、ほとんどどのような相関係数でも得られる!

地区数	相関係数の 最小値	相関係数の 最大値
6	-0.99	0.99
12	-0.99	0.99
18	-0.97	0.99
24	-0.92	0.99
30	-0.73	0.98
36	-0.71	0.96
42	-0.55	0.95
48	-0.50	0.90
54	-0.42	0.82



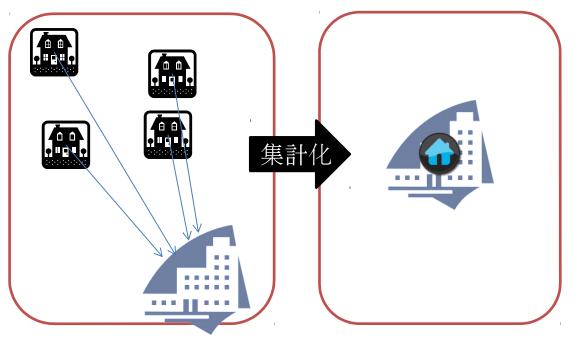
地区数 12 で r = 0.87 となるゾーニング例

ゲリマンダー 空間単位の恣意性



位置情報の集計化問題

- ■地点間の距離の変数を利用する施設配置モデルや空間選択モデルでは、距離を 定義する位置座標が必要である。
- ■しかし、個人の住所や目的値は地理的集計単位でしか得られない場合が多い。
- ■その場合、各単位内部の位置座標は、重心などの代表点座標で代替され、計測 される距離に集計によるバイアスがうまれる。



極端な例だが、各住居から施設までの距離は、地理的集計によって、施設の位置も全ての住居も、地区の代表点での住居も、地区の代表点であるとみなどがある場合、施設までの移動に必要な距離

はゼロになってしまう。

地理情報科学教育用スライド © 中谷友樹

どうすればよいのか

- 解釈や分析方法に注意する
 - -集計された情報しかないならば、MAUPの一般的な解決方法はない
 - -分析単位によって結果が変わりうる点に留意し、分析単位のスケールに対応する解釈を行う。
- 適切な分析単位を作成する
 - -統一した基準で地理的単位を作成する(ゾーンデザイン)。
 - -集計データから非集計データを推計するマイクロシミュレーションや、空間的補間 による位置情報の詳細を推計する
- 空間的集計による影響をモデル化する
 - -集計によるバイアスそのものを説明・予測するモデルを考える
 - •空間的自己相関と相関係数のスケール変化
 - 位置情報や距離などの幾何学的な集計問題など

ローカルパターンと グローバルパターン Local & global patterns

グローバルな空間分析の問題 空間的集積性の例

- ■空間的なパターンを調べるために、Moran's Iのような空間的自己相関の統計量を用いることが多い。
- ■正の空間的自己相関は、全体として地理的な集積傾向を示す(グローバル・パターンの指標)。
 - ■/>E[/] = 1/(n-1) であれば、正の空間的自己相関
- ■しかし、正の空間的自己相関が有意であったとして、具体的にどこに集積しているのか(ローカル・パターン)については、分からない。

Moran's I (自己相関係数)

$$W = \sum_{i}^{n} \sum_{j}^{n} w_{ij}$$

LISA (Local Indicators of Spatial Association) Local Moran's I

- \blacksquare Moran's I は、各地区 I の自己相関成分 I_i の総和で定義されている。
- \blacksquare この I_i (ローカルモラン)に、各地区 i を中心とする自己相関傾向のローカルな情報が含まれている。

$$I = \frac{1}{W} \cdot \frac{\sum_{i}^{n} (x_{i} - \overline{x}) \sum_{j}^{n} w_{ij} (x_{j} - \overline{x})}{\sigma^{2}}$$

$$= \sum_{i}^{n} I_{i}$$

$$I = \frac{1}{n} \cdot \frac{\sum_{i}^{n} (x_{i} - \overline{x}) \cdot \left(\sum_{j}^{n} w_{ij} x_{j} - \overline{x}\right)}{\sigma^{2}}$$

$$= \sum_{i}^{n} I_{i}$$

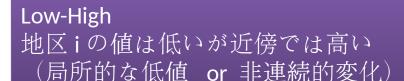
$$= \frac{1}{W} \cdot \frac{(x_{i} - \overline{x}) \sum_{j}^{n} w_{ij} (x_{j} - \overline{x})}{\sigma^{2}}$$

$$= \sum_{i}^{n} I_{i}$$

$$I_{i} = \frac{1}{n} \cdot \frac{(x_{i} - \overline{x}) \cdot \left(\sum_{j}^{n} w_{ij} x_{j} - \overline{x}\right)}{\sigma^{2}}$$

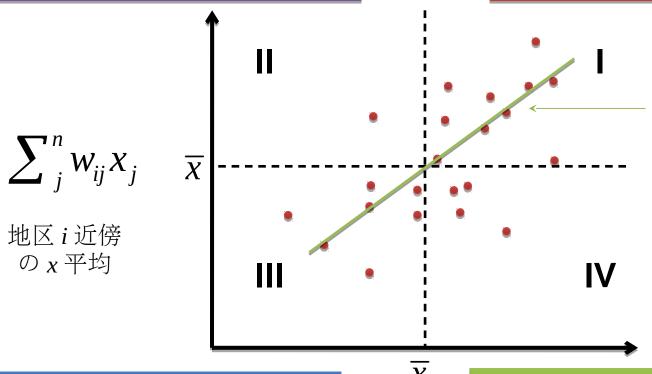
$$I_{i} = \frac{1}{n} \cdot \frac{(x_{i} - \overline{x}) \cdot \left(\sum_{j}^{n} w_{ij} x_{j} - \overline{x}\right)}{\sigma^{2}}$$

モラン散布図 ローカルモランの解釈



High-High

地区iとその近傍がともに高い (高い値のローカルな集積)



回帰直線 (この傾きが Moran's I に相 当)

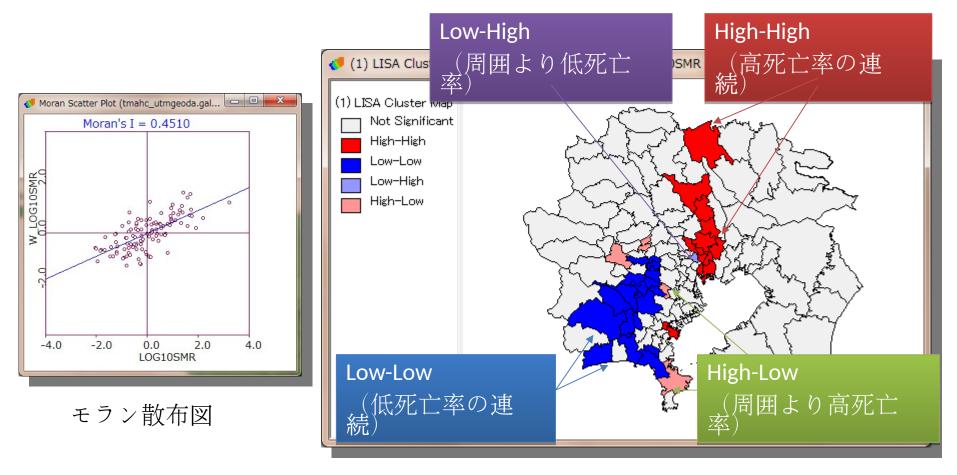
Low-Low

地区iとその近傍がともに低い (低い値のローカルな集積) $ar{X} \ m{X}_i$ 地区iの

High-Low

地区iの値は高いが近傍では低い (局所的な高値 or 非連続的変化)

ローカルモランの事例



5% 水準で有意なローカルモランの係数値を、モラン散布図で理解できる4 タイプ別に 地図化した。高い死亡率のセクターと、低い死亡率のセクターの地理的な集積が明らか 東京大都市圏保健所管轄区別標準化死亡比(1990年)

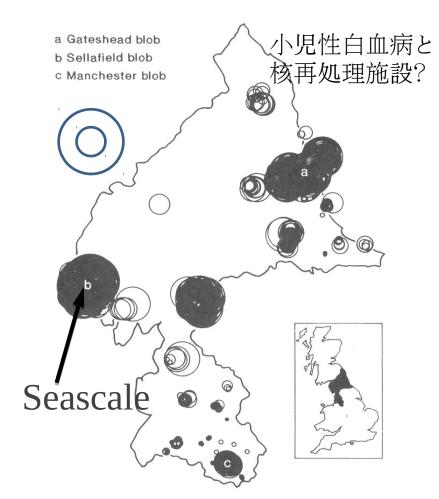
グローバルとローカルな集積

- ■グローバルな集積性の分析
 - 全体として集積する傾向の識別 Clustering tendency
 - 結果は1つの数値に要約
- ■ローカルな集積性の分析
 - 具体的に集積する場所の特定 Cluster detection
 - 結果は地図に示しうる
 - ただし、ローカルな集積の範囲をどの程度に想定する のかにおいて、ある種のスケール問題が発生する

Geographical Analysis Machine

ローカルな関係をみるスケール問題への挑戦

- 1. 適当な位置・大きさの円を 仮定する
- 2. 円内部の有病率を統計的に 検定し、有意に高いなら円 を描く
- 3. 円の位置、大きさを変更し 1と2を「かたっぱしに」 繰り返す

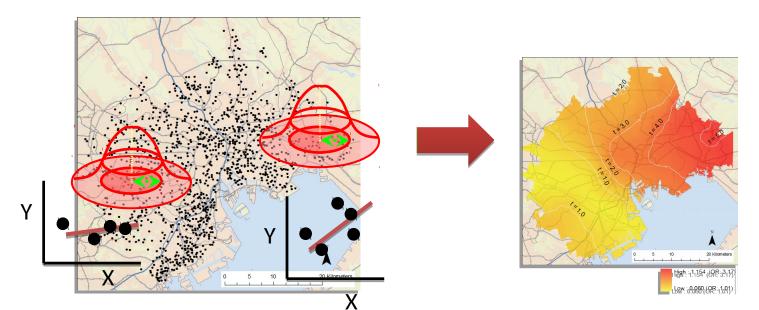


Source: Openshaw, S. et al 中等方式科学教育用スライド ©中谷友樹

地理的加重回帰法 GWR 回帰分析のローカルモデリング

- 回帰係数を対象地域全体で1つに固定する通常の回帰モデ $Y_i=eta_0+\sum_k^Keta_k x_{k,i}+arepsilon_i$ ■ グローバル・モデル: ル
- ローカル・モデル: 回帰係数のローカルな変化を想定したモデル(GWRモデル)

$$oxed{u}$$
ローカルな係数は、地区 i 近傍のデータに重みをつけた最小 2 乗法によって推定する。 $Y_i=eta_0(u_i,v_i)+\sum_k^Keta_k(u_i,v_i)x_{k,i}+arepsilon_i$



地点ごとに回帰 係数の推定を繰 り返すと、回帰 係数の分布面が 得られる

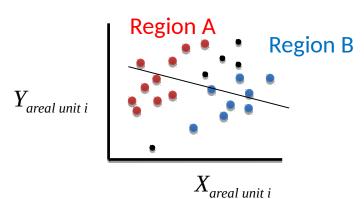
地理情報科学教育用スライド © 中谷友樹

マルチレベル分析

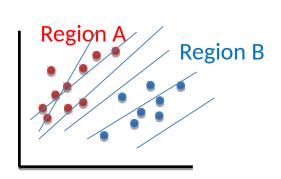
特定の地域グループを前提として、パラメターのグループ間での変動を 、適当な確率分布でモデル化する階層線型モデル(マルチレベル分析) も、ローカルな回帰モデルの1種とみなせる。

分析単位レベルのモデル
$$Y_i = \beta_{0(j)} + \sum_k^K \beta_{k(j)} x_{k,i} + \varepsilon_i$$

グループ・レベルのモデル
$$egin{bmatrix} eta_{0(j)} \sim N(oldsymbol{\gamma}_0, \sigma_0^2) \\ eta_{k(j)} \sim N(oldsymbol{\gamma}_k, \sigma_k^2) \end{bmatrix}$$
 j は地区 i が属している地域グループ



グローバルな回帰モデル 切片と傾きは全体で1つ



マルチレベル・モデル 切片と傾きは地域グループごと変動

まとめ

- ■空間分析における空間スケールの 2 つの問題
 - ◆可変地区単位問題 MAUP
 - ◆分析単位の問題
 - ◆空間分析に基づく推論上の誤りに対する注意
 - ◆解決の難しい難問
 - ◆適切な空間「単位」とは?
 - ◆ローカルパターンとグローバルパターン
 - ◆分析範囲の問題
 - ◆グローバル: 全体の傾向の要約
 - ◆ローカル: 特定の位置・範囲に関する特性(地図上に示せる)
 - ◆適切なローカルな範囲とは?