

CSIS Discussion Paper Series

No. 11

二つの領域分割図の位相的適合状態を
評価する方法と統合化する方法

岡部篤行 貞広幸雄

April 7, 1999

東京大学 空間情報科学研究センター
113-8656 東京都 文京区 本郷 7-3-1

1 はじめに

ある地域における属性分布のデータは、その地域を分割した地区ごとの属性データとして与えられる場合が多い。例えば、小金井市の人口分布が、町丁目単位の人口で与えられるといった場合である。このような地域分割した地区の集まりを「領域分割図」と呼ぼう。

領域分割図は、いろいろな主体によって作られることが多い。例えば、小金井市の町丁目単位の領域分割図は、統計局でも国土地理院でも作成している。小金井市の町丁目属性データを分析する時、統計局作成の属性データを統計局の作成した領域分割図で分析し、一方で国土地理院作成の属性データを国土地理院の作成した領域分割図で分析する限り、二つの領域分割図をつき合わせる必要がないので、問題は生じない。問題は、統計局の属性データと国土地理院の属性データを複合的に利用して分析をしたい場合に生じてくる。同じ地域の同じ領域分割図であるから、どちらも同じであるのが期待されるが、作成した主体が違うので、それは保証の限りでない。むしろ違う方が多いと言えよう。

このことを考えると、地域の分析を複数の主体がそれぞれに作った領域分割図単位のデータで行おうとすると、まずは分析に入る前に、複数の領域分割図がどの程度一致しているかを調べる必要が出てくる。もしも一致の度合いが悪いとなると、いくら精緻な分析をしても意味がないと言える。そこで、一致していない場合には、複数の領域分割図を一つに統合することが必要となってくる。これらの問題は、今までに、いくつかの研究がなされている。しかし、汎用的な方法を提案している研究は少ない。そこで、この論文では、この二つの問題に関し、「位相的問題」と「計量的問題」に分けて、前者の問題について汎用的な方法を提案しよう。

Figure 1: 二つの領域分割図 $\mathcal{S}_1, \mathcal{S}_2$

Table 1: 名称リスト M_1, M_2

M_1 の名称		M_2 の名称	
1	緑町	1	桜町
2	桜町	2	緑町
3	本町	3	本町
4	貫井町	4	貫井町
5	梶野町	5	梶野町
		6	前原町

2 問題

以上の問題を、もう少し明確に述べておこう。いまある地域 S があるとし、その地域を領域分割した領域分割図が二つ、 $\mathcal{S}_1 = \{S_{11}, \dots, S_{1n_1}\}$, $\mathcal{S}_2 = \{S_{21}, \dots, S_{2n_2}\}$ があるとしよう。すなわち領域分割図 \mathcal{S}_1 は地区 S_{11}, \dots, S_{1n_1} より成り立ち、領域分割図 \mathcal{S}_2 は地区 S_{21}, \dots, S_{2n_2} より成り立っているとする (図 1 参照)。また同時にこれらの地区の名称のリスト $M_1 = \{m_{11}, \dots, m_{1k_1}\}$ と $M_2 = \{m_{21}, \dots, m_{2k_2}\}$ が与えられているとしよう (表 1)。 (\mathcal{S}_i, M_i) の対を「名称付き領域分割図」と呼ぶことにする。

そうすると、問題は以下ようになる。

問題 1 名称付き領域分割図 (S_1, M_1) と (S_2, M_2) の適合度合いを評価する方法を考案せよ。

問題 2 名称付き領域分割図 (S_1, M_1) と (S_2, M_2) を統合して一つの領域分割図にする方法を考案せよ。

当論文では、問題 1、中でも特に位相的な適合度合いについて検討することにする。

3 二つの名称付き領域分割図の適合度合いの評価方法

3.1 名称リストの対応評価

まず、最初に行うべき評価は、名称リストの対応についてである。表 1 のようにそもそも町の数合わないことは、現実には往々にしてあることである。そこで、最初のチェックは名称リスト M_1 と M_2 に関して、 $m_{k_1} = m_{k_2}$ を確かめることである。

もし、これが成り立てば、次の作業に進む。そうでない場合は、入力の過ちが疑われ、修正を行うことになる。例えば表 1 の場合、 $m_{k_1} = 5$, $m_{k_2} = 6$ であるから、表 2 のように修正する。

二つの名称リストに含まれる地区の数が一致しても、異なる地区があがっているかもしれない。そこで次は、名称リストの名称が 1 対 1 に対応しているかのチェックが必要である。表 2 の場合、 M_1 には梶野町があり M_2 にはない。一方、 M_2 には前原町があり M_1 にはない。これはそもそも入力のミスが疑われ、修正を行うことになる。例えば表 3 のような修正を行う。

二つの名称リストの地区の名称が 1 対 1 に対応して、初めてつぎの作業に進むことができる。

Table 2: 名称リスト M_1, M_2 の修正 1

M_1 の名称		M_2 の名称	
1	緑町	1	桜町
2	桜町	2	緑町
3	本町	3	本町
4	貫井町	4	貫井町
5	梶野町	5	前原町

Table 3: 名称リスト M_1, M_2 の修正 2

M_1 の名称		M_2 の名称	
1	緑町	1	桜町
2	桜町	2	緑町
3	本町	3	本町
4	貫井町	4	貫井町
5	前原町	5	前原町

Figure 2: データベースの領域分割の修正

以上、述べてきた名称リストの一致は、当然、満たされていると期待することではあるが、地区の数が多いと、往々にして起きる不一致で、最初に行うべき作業であろう。なお、この作業は地図を使わずに出来る簡単な作業である。

3.2 領域分割図データが分割図となっていることの検定

領域分割図とは数学的に言うと、

$$S = S_{i1} \cup \dots \cup S_{in_i}, |S_{ij} \cap S_{il}| = 0, j \neq l, j, l = 1, \dots, n_i (i = 1, 2) \quad (1)$$

が成り立つことである。一方、領域分割をコンピュータで扱うには、それを数値データで表現しなければならない。その場合、数値データで表現された領域分割図が、数学的に領域分割となっているという保証はない。例えば、図 2(a) などは、よく見ると、線分がつながっていないで領域分割とはなっていない。そこで図 2(a) のように修正を行う。この作業は、もっぱら領域分割をどのようなデータ構造で扱うかによる (GIS ソフトによる) winged-edge のデータ構造であれば、領域分割をチェックする算法が提案されている。

Figure 3: 名称の地区と領域分割図の地区が1対1でない場合(1)

3.3 地区名称と領域分割図の地区との対応付け

領域分割データが領域分割となっているとなると、次には、地区名称と地区の対応付けを行う必要がある。まず、名称リスト M_i にある地区の数 m_k と領域分割図 S_i にある地区の数 n_i が合致していなければならない ($i = 1, 2$)、 $n_i = m_k$ であれば次の作業に進むが、そうでない場合は、間違いが生じているので修正を行う必要がある。

次に、地区名称と領域分割図の地区との対応が1対1になっているかのチェックを行う。まず $n_i = m_k$ であっても、図3(a)のような場合がありうる。この場合は1対1になるように修正を行う図3(b)。

次に $n_i \neq m_k$ の場合、二つの場合に分ける必要がある。まず $n_i < m_k$ の場合である。これは例えば図4(a)のように、同じ名称の地区が飛び地している場合におこることである。このように地区が分かれてしまっている場合には、図4(b)のように新たに地区の名称を付け、名称リストを表4のように修正する。

次に $n_i > m_k$ の場合である。これは例えば図5(a)のように、地区の区分分けが十分でなかった場合である。この場合は、図5(b)のように地区を細分する修正を行う。

Figure 4: 名称の地区と領域分割図の地区が 1 対 1 でない場合 (2)

Table 4: 名称リストの修正 3

	M_1 の名称
1	緑町
2	桜町
3	本町 1
4	貫井町
5	本町 2
6	前原町

Figure 5: 名称の地区と領域分割図の地区が 1 対 1 でない場合 (3)

以上の作業が済むと名称リストの地区と領域分割図の地区が 1 対 1 に対応することになり、ここで初めて、二つの領域分割図の適合性を検討することができる。このような状態になった名称付き領域分割図 (S_1, M_1) と (S_2, M_2) を「相互比較可能な領域分割図」と呼ぶことにする。

3.4 相互比較可能な領域分割図の位相構造化

もちろんのことながら、二つの名称付き領域分割図 (S_1, M_1) と (S_2, M_2) が相互比較可能な領域分割図であるということは、二つの領域分割図が同じであるということではない。例えば、図 6 の二つの領域分割図は相互比較可能であるが、領域分割図として同じでないのは明らかである。

相互比較可能な領域分割図が同じというためには、まず、地区の「並び」が同じである必要がある。ここで「並び」とは、緑町の隣には桜町、貫井町、本町がある、といった位相関係が同じということである。この位相関係を捉えるために、領域分割図をグラフで表すことにする。その手順はまず、領域分割図 S_i の孤立ループを抽出し、それぞれのループを 1 つのノードと一つのリンクで表す (図 7)。次に、孤立ループ以外は、度数が 3 以上の頂点 (黒丸) をノードとし、それらのノードと連結するノードをリンクとして表す (図 7)。

Figure 6: 相互比較可能な領域分割図

Figure 7: 領域分割図のグラフ化

その結果できるノードとリンクの集合の対は地図グラフとなり、この地図グラフを G_i で表して、 S_i の「領域分割地図グラフ」と呼ぶことにする。

一般に地図グラフ G_i は双対グラフをもっている。そのグラフを G_i^* で表すことにし、「領域分割双対地図グラフ」と呼ぶことにしよう(図8参照)。なお領域分割双対地図グラフでリンクは、地区の隣接関係を示している。

3.5 領域分割地図グラフの位相適合度の評価

地図グラフができると、位相関係が合っているのかを調べることができる。図6の場合でそれを見てみよう。図6の領域分割双対グラフはそれぞれ図9

Figure 8: 領域分割グラフと領域分割双対グラフ

Figure 9: 二つの領域分割双対グラフ

の (a)(b) のようになる。二つの地区どうしの隣接関係だけを見て行くと、破線の部分で二つの隣接関係が異なっていることが分かる。全体的に評価するには、例えば

$$T_1 = \frac{\text{隣接関係の同じリンクの数}}{\text{全てのリンクの数}} \quad (2)$$

なる指標が考えられる。図 9 の場合 $T_1 = 0.75$ となる。

もう少し複雑な関係を考慮する二つの評価方法を示そう。まず二つの地区の位相的な関係は、図 10 に示したような関係に分けることができる。領域分割 S_i における二つの地区の間の距離 $d(\{S_{1j}, S_{1k}\}, \{S_{2j}, S_{2k}\})$ を図 10 のように定める。すなわち S_{1j} と S_{1k} の関係を示す図から S_{2j} と S_{2k} の関係を示す図までの最短経路で幾つ位相関係が変わるかで、距離を測る。そうすると領域分割図 S_1 と S_2 における全体の平均的な位相的ずれ $D(S_1, S_2)$ は、

$$D(S_1, S_2) = \frac{\sum_{i,j=1,\dots,n} d(\{S_{1j}, S_{1k}\}, \{S_{2j}, S_{2k}\})}{\text{全てのリンク数}} \quad (3)$$

で評価することができる。

以上二つの指標は、地図グラフである性質を無視している。それを考慮した指標を示そう。領域分割 S_i の地区 j の隣接関係を反時計回りの輪環リスト L_{ij} をつくる。例えば、図 9(a) の桜町の場合、 $L_{1j} = (\text{緑町、貫井町、前原町、本町})$ となり、(b) の場合 $L_{2j} = (\text{貫井町、緑町、前原町})$ となる。これらの輪環リストは局所近傍位相を示したものである。この二つの輪環リストが一致すると、局所近傍位相は同相であると言える。この概念を使って二つの領域分割図の局所近傍位相適合性 N を

$$N = \frac{L_{1j} = L_{2j} \text{ であるノードの数}}{\text{全てのノードの数}} \quad (4)$$

Figure 10: 二つの地区の位相関係とその間の位相距離

と指標化することができる。

3.6 無名称化した位相の適合性

さて以上での位相チェックで位相構造が一致していないことが分かった場合、次の手続きは、二つの方針がある。一つは、位相構造のずれはずれとしておき、計量的に二つの領域分割図のずれを評価する方法である。もう一つの方針は、あくまでも位相構造を修正してから、計量的なずれの評価をする方法である。前者を「計量優先型」、後者を「位相優先型」と呼ぶことにする。

以下では、位相優先型について、手続きを検討してゆくことにしよう。まず位相構造の不一致である理由は二つが考えられる。まず、そもそも(無名称化した)位相構造が異なっている場合である。次に位相構造が同じにもかかわらず、単純に名称を振り間違えたという場合である。いずれにせよ、位相構造が同じかどうかを調べる必要がある。

図9で示した二つの領域分割図の場合、ノード名称付きの地図グラフが一致しなかった。しかしノードの名称を取ると、二つの地図グラフは同相である。名称をとった地図グラフを「無名称地図グラフ」と呼ぶことにしよう。さて問題は

問題3 二つの無名称地図グラフが同相であるかどうかを調べる方法を考案せよ。

となる。一般に平面グラフの同相問題は、多項式オーダの計算時間で解く coding 法が提案されている。ここではグラフが地図グラフでかつ領域分割図であるという特殊性を利用して、それに変わる新たな方法を提案しよう。

まず、 S_1 の双対領域分割グラフの外部境界に接する地区を抽出する(図11(a)の細線の部分)。この抽出した部分は、環状 R_{11} になっている。同様に、 S_2 の双対領域分割グラフの外部境界に接する環状地区 R_{12} を抽出する。

Figure 11: 皮むき法

次に、 R_{11} を固定して R_{12} を回転させ、 R_{11} と R_{12} が一致するかどうかを調べる。もし一致するところがなかったら、同相ではありえない。

一致した場合、 R_{11} と R_{12} を除去する (図 11(b))。それに対して上と同じ作業を行い R_{21} と R_{22} を求める (図 11(c) の細線の部分)。 R_{21} と R_{22} が一致するかどうかを調べる。もし一致するところがなかったら、同相ではありえない。

以下、同様にして「皮」をむくようにして、全てが同相であれば、無名称領域分割地図グラフは同相である。

Figure 12: 位相の違いが起きやすい場合

3.7 無名称地図グラフが同相である場合の名称振り間違いの検出

無名称地図グラフが同相であるとなると、名称の振り間違いである可能性がある。例えば図9の場合、その可能性がある。この場合、どちらかが正しいものであるというのが分かっているならば、修正は簡単で、正しい方の名称リストと地区対応表に合わせればすむ。

どちらが正しいか分からない場合は、所詮、この段階で直しようがない。再度、最初に戻ってどこが違っているかを再検討するしかない。この段階でできることは、輪環リストが違うところを図示して、修正を容易にすることであろう。

3.8 局所近傍位相の修正

無名称地図グラフが同相でないとなると、位相構造が異なっていることになる。この場合、問題は複雑である。名称の振り間違いに加え、位相構造の違いが重なるからである。であるから方法は、ヒューリスティックとならざるをえない。その一つを以下に示そう。

位相構造の違いは、二つのノードが近い場合に起きやすい。例えば図12

Figure 13: 図 12 の双対領域分割グラフ

を参照されたい。このような領域分割図の場合、デジタイズの具合によって、両方、ありえることである。しかし位相的には図 13 のようになり、同相ではない。であるが、ノードの付け替えで、同相になりうる状態である。このような場合は、修正してどちらかに統一するのがよいであろう。

この修正を行う手順は次のようにすればよい。

- Step 0 入力データ：ノードの輪環リスト、領域分割データ、閾値距離
- Step 1 領域分割地図グラフのノード間距離が閾値距離より小さいリンクを選ぶ。
- Step 2 そのリンク l_j に対応する双対領域分割地図グラフのリンク l'_j を定める (図 14(a) 参照)
- Step 3 リンク l'_j を共有する双対領域分割図グラフの地区をなすノードの内、リンク l'_j の端点ノード以外の二つのノードを求める (図 14(b) の太線の円)
- Step 4 リンク l'_j をその二つのノード間のリンク l''_j にフリップする (図 14(c))

Figure 14: 局所位相の修正

Step 5 リンク l_j'' に対応するように領域分割地図グラフの位相を修正する(図 12(b)の領域分割となる)

4 おわりに

以上の位相修正によって最終的に位相構造が一致する場合もあれば、そうでない場合もある。ない場合には、最初に戻って出直すこととなる。一致した場合を「相互計量比較可能な領域分割」と呼ぶことにし、その評価方法は、稿をあらためて考案することにする。

謝辞この研究は、科学技術事業団「空間情報科学研究のためのデータベース開発研究」で行われている研究の一部である。研究会の方々の貴重なコメント、特に奥貫圭一氏のコメントに感謝する。