空間相関関数とその統計的検定の実用的計算手法と視覚化

実用的メッシュデータ解析システム構築のための空間相関分析法の体系化 その1

PRACTICAL CALCULATING METHOD AND VISUALIZATION OF SPACE
CORRELATION AND ITS STATISTICAL TEST

Synthesis of Urban Lattice Data Analysis for Practical Use, Part 1

青木 義次*, 大佛 俊泰**

Yoshitsugu AOKI and Toshihiro OSARAGI

The space correlation analysis (SCA) was proposed with its possibility of application on urban analysis in our previous papers. The calculating program of SCA, however, is not easy to develop for urban planners because the program has a relatively large size and complicated. The calculating theory is reconstructed to make SCA fit for practical use. The following possibilities of calculating method of SCA are shown; 1) the space influence function (SIF) can be calculated by using ordinary library program, 2) the space correlation function (SCF) can be calculated very rapidly by using FFT (fast Fourier transform).

Also presented in this paper are visual representation methods of SCA model with reliable range of estimated parameters. It is very important in actual planning to know the structure of spatial influence relation of urban activities accurately and intuitionally. Using the statistical range estimation and testing theory, 1) the method for getting the confidence interval of SCF is shown, and 2) the method for grouping the value of SCF is shown. The visual presentation of the results of the above methods are very effective to assist urban planners to understand the characteristics of spatial relation of urban activities.

Keywords : space correlation function, statistical test, visualization, library program, urban activity, FFT

1. はじめに

近年の都市計画・地域計画においては現状の都市・地域の状態を把握するためにメッシュデータを用いて空間的に分布している様々な特性を定量的に把握することが試みられており、また、このための解析方法に関する研究もなされている。これまでに筆者の一人は、メッシュデータが持っている空間的な配置情報を有効にいかすための解析方法として空間相関分析法 (SCA 法) の基礎的理論を報告し¹⁾、また、この応用として、土地利用における連担性、共存性、排斥性の概念も空間相関関数により定式化可能で計量的解析ができることを示した²⁾。さらに、空間相関関数を用いて都市活動の空間的相互影響関係を空間影響関数モデルとして抽出してゆく方法を提案し、通常回帰モデルと比較して SCA 法の有効性を示した³⁾。

しかし、この SCA 法を実際の都市計画立案等の場で

実用的に用いるためには、以下のような具体的な問題がある。すなわち、

- ① SCA 法の計算を誰にでも容易にできること、
- ② SCA 法の解析結果から、計画者 (利用者) が、都市活動要素の空間的影響関係を容易にイメージできること、
- ③ 各パラメータの推定においても、推定に使用したデータが少ないときには、数回のコイン投げで推定したコインの表の確率が不正確であるように、推定した値の前後にふれている可能性がある。このような推定値のふれている可能性を SCA 法を利用する計画者が判断できることが望ましいこと、
- ④ 推定結果が、単に統計的に優れたモデルというだけでなく、推定パラメータの意味を計画者が判断するときに誤りの生じないような推定になっていること、
- ⑤ 場所ごとの特殊な空間構造にも柔軟に対応できるこ

本論文は、文献4)で報告したものを修正・加筆したものである。

* 東京工業大学 助教授・工博

** 東京工業大学 助手

Assoc. Prof. of Tokyo Institute of Technology, Dr. Eng.
Research Associate of Tokyo Institute of Technology

と、

などの課題を解決する必要があると言える。

①については、これまでの SCA 法が若干複雑な計算を必要とし、その計算プログラム作成が容易でないことから、第2章で、

1) SCA 法の理論体系を再検討することで、既存のソフトウェアを利用し高速に計算する方法を開発する。

②および③については、第3章で、

2) 都市活動要素の空間的影響関係を、視覚的にディスプレイするとともに、推定値の有効範囲をもディスプレイする方法を提案する。

④については、上記の視覚化のほかに、推定結果が誤差最小化という統計的基準だけでなく、計画的に有効な知見を誘導するような視点からのパラメータ安定推定の方法を、⑤については、SCA 法に場所性の概念を導入したモデル化の方法を、次報で報告したい。

2. SCA 法における計算方法の再構成

2.1 空間相関関数とメッシュデータのとり方

SCA 法の計算手続きを既存のソフトウェアを利用して実行するために、これまでの理論体系を等価変形し、既存の線型連立方程式、重回帰分析、スペクトル解析の理論枠組みとの関係を明示できるようにしておく。

都市空間を二次元直交座標系（メッシュ空間）で表し、対象都市領域 M の各メッシュ (t_1, t_2) において n 個の変数 X_i ($i=1, 2, \dots, n$) の値 $x_i(t_1, t_2)$ が得られているものとする。変数 X_i の空間平均 $\bar{x}_i$ は次のように定義されるものであるとする。

$$\bar{x}_i = \lim_{T_1 \rightarrow \infty} \lim_{T_2 \rightarrow \infty} \frac{1}{T_1 T_2} \int_{-T_1/2}^{T_1/2} \int_{-T_2/2}^{T_2/2} x_i(t_1, t_2) dt_1 dt_2 \quad \dots\dots\dots (1)$$

さらに、各地点からメッシュで東へ k_1 個、南へ k_2 個ずれた場所との関係を調べるために、この空間的な『ずれ』 (k_1, k_2) のことを『空間ラグ』と呼び、この空間ラグごとに、すなわち、ある地点での変数 X_i と (k_1, k_2) だけずれた場所での変数 X_j との類似性を測るための量として、(相互) 空間共分散関数 $C_{ij}(k_1, k_2)$ を以下のように定義する。

$$C_{ij}(k_1, k_2) = \lim_{T_1 \rightarrow \infty} \lim_{T_2 \rightarrow \infty} \frac{1}{T_1 T_2} \int_{-T_1/2}^{T_1/2} \int_{-T_2/2}^{T_2/2} x_i(t_1, t_2) \cdot x_j(t_1 + k_1, t_2 + k_2) dt_1 dt_2 \quad \dots\dots\dots (2)$$

この量は、普通に用いられる共分散の概念を一般化したものである。また、相関係数が共分散を分散で基準化して得られるように、(相互) 空間相関関数は以下の定義から計算できる。

$$R_{ij}(k_1, k_2) = C_{ij}(k_1, k_2) / \sqrt{C_{ii}(0, 0) C_{jj}(0, 0)} \quad \dots\dots\dots (3)$$

この量は (k_1, k_2) ずれた場所との相関係数と考えられる

図-1 空間相関関数を計算するときのデータのとり方

量になっている。

ここで、対象都市領域 M 内の任意の座標 (t_1, t_2) において各変数 X_i の値が得られているとき、そこから空間ラグ (k_1, k_2) だけ離れた座標 $(t_1 + k_1, t_2 + k_2)$ とのペアをとり空間ラグ (k_1, k_2) に関する空間共分散関数を計算する場合のデータのとり方には次の2つの方法がある。例えば、領域 M の大きさを $N_1 \times N_2$ とし、空間ラグが $(-L_1 \leq k_1 \leq L_1, -L_2 \leq k_2 \leq L_2; L_1=3, L_2=3)$ の範囲について空間共分散関数を計算する場合、

(a) 図-1(A) に示すように、空間共分散関数は領域 M 内で空間ラグが (k_1, k_2) となる可能なデータのペアのすべてから計算されるが、データの組合せ数は空間ラグ (k_1, k_2) によって異なり $(N_1 - |k_1|)(N_2 - |k_2|)$ 組となる。以下では、このようなペアのうち座標 (t_1, t_2) のとり得る範囲を $M(k_1, k_2)$ と表記する。

(b) 図-1(B) に示すように、上記と同じデータに対して、空間共分散関数を計算するペアの数を空間ラグ (k_1, k_2) によらず一定数の $(N_1 - 2L_1)(N_2 - 2L_2)$ 組となるように計算することも可能である。この場合は、すべての k_1, k_2 について上記の $M(k_1, k_2)$ の共通集合部分（以下 M_0 と表記する）に座標 (t_1, t_2) を限定したことに等しい。したがって、同じデータに

対して上記の方法に比べ周辺部の座標におけるデータを利用しないことになる。

つまり、前者はメッシュデータの端部の情報を失わないという利点を持ち、後者は各空間ラグにおけるデータ数は一定であるという利点を持つ。既に報告したように^{注2)}空間共分散関数またはそれを規準化した空間相関関数を求めて、都市活動の連担性、共存性、排斥性を計量的に分析する際には、(a)の方法が精度上望ましい。しかし、実用的な見地からすると、次のような条件下では、(b)の方法で代替できる。

- ① 対象領域が十分広く (b)の方法で捨象する周辺部データの数が相対的に小さくなる場合
- ② 分析対象地域の空間構造の特徴を見るとき、対象地域周辺では外部地域の影響のためその地域の特徴がむしろ曖昧になっており、周辺部のデータよりも中心部のデータを活用しようとする場合

また、このような代替によって精度が落ちることの対策として、後述の推定値の信頼区間を調べる方法によって、どれだけ精度が落ちたのかが分かる。経験的には、空間構造すなわち都市活動間の影響の仕方が対象地域中央部と周辺地域で異なるような場合には、この差が大きく、中央部と周辺部でその影響の差がない場合には、推定値そのものは大きくずれることはないが、データ数を落とした分だけ推定値の信頼区間は広がってしまう^{注4)}。

2.2 空間影響関数モデルと推定方法

ある地点の変量 X_i が、その付近の他の変量 X_i ($i=1, 2, \dots, n$) の影響を受けており、また、その影響の程度は地点によらず空間ラグ ($l_1, l_2; -L_1 \leq l_1 \leq L_1, -L_2 \leq l_2 \leq L_2$) だけに依存しているものと考え、つぎの空間影響関数モデルを定式化することができる^{注1)}。

$$x_j(t_1, t_2) = a_0 + \sum_{i=1}^n \sum_{l_1=-L_1}^{L_1} \sum_{l_2=-L_2}^{L_2} h_{ij}(l_1, l_2) \cdot x_i(t_1 - l_1, t_2 - l_2) \dots \dots \dots (4)$$

このとき、(4)式を(2)式の空間共分散関数の定義式に代入することにより、空間影響関数 $h_{ij}(l_1, l_2)$ と空間共分散関数 $C_{ij}(k_1, k_2)$ との関係式として、下記の(5)式を得ることができる。ここで、 C_{ij} 、 C_{ii} の値が既に計算されているならば、未知数である $h_{ij}(l_1, l_2)$ の数は合計 $n(2L_1+1)(2L_2+1)$ 個であり、方程式の数も $n(2L_1+1)(2L_2+1)$ であることから、一般には解くことができ、 $h_{ij}(l_1, l_2)$ は連立方程式(5)の解として求めることができる。さらに、モデルの未知定数項 a_0 も、得られた空間影響関数 $h_{ij}(l_1, l_2)$ を用いることで、(6)式により求めることができる。

$$C_{ij}(k_1, k_2) = \sum_{k=1}^n \sum_{l_1=-L_1}^{L_1} \sum_{l_2=-L_2}^{L_2} h_{ik}(l_1, l_2) \cdot C_{ik}(t_1 - l_1, t_2 - l_2) \dots \dots \dots (5)$$

図—2 SCA法における空間影響関数の推定過程

$$a_0 = \bar{x}_j - \sum_{i=1}^n \left[\bar{x}_i \sum_{l_1=-L_1}^{L_1} \sum_{l_2=-L_2}^{L_2} h_{ij}(l_1, l_2) \right] \dots \dots \dots (6)$$

図—2は、モデルの定式化から空間影響関数の推定に至るまでのSCA法の計算過程を簡単にまとめたものである。SCA法の計算プログラムは自己開発するには若干大規模なものであるが、この空間影響関数モデルを推定する(5)、(6)式の解法に既存の線型連立方程式解法のパッケージ・プログラムを利用すれば、プログラム開発に要する労力の削減を図ることができる。

2.3 重回帰分析パッケージ・プログラムの利用

前項では、既に計算された空間共分散関数をもとに、空間影響関数モデルのパラメータを推定する際に、線型連立方程式のパッケージ・プログラムを活用できることを示したが、この方法は、前述の活用データの取り方としては、(a)、(b)のどちらでもかまわない。

以下では、活用データの取り方で、(b)の方法でも構わないと判断できる場合に重回帰分析パッケージ・プログラムを活用して、空間共分散関数の計算自体も既存プログラム内で計算させ、原メッシュデータから空間影響関数モデルを直接推定する方法を提示したい。

分析対象となる対象都市領域 M の大きさを $N_1 \times N_2$ とするとき、都市活動を表すひとつひとつの変量 X_i 自体のメッシュデータはいわば位置座標を表す2つのサフィクスをもつ二次元配列になっており、空間的位置関係の情報を有している。一方、重回帰分析モデルでは、ひとつの変量 X_i は、例えば何番目ということを表す1つのサフィクスをもつ一次元配列データの形をしている。この位置情報を損なうことなく、各変量の二次元配列を一次元化して、重回帰分析プログラムのデータとして与えなくてはならない。このデータの変換を機械的に行うため、次の1対1写像 $P(i, l_1, l_2)$ によって(4)式で定式化された空間影響関数モデルの各変量の値に並び替えを施す。

$$P(i, l_1, l_2) = (i-1)(2L_1+1)(2L_2+1)$$

$$+(l_1+L_1)(2L_2+1)+l_2+L_2+1 \\ \dots\dots\dots(7)$$

for $i=1, 2, \dots, n$ (n : 変数の数)

$$-L_1 \leq l_1 \leq L_1, \quad -L_2 \leq l_2 \leq L_2$$

ここで並び替えを行った後の説明変量を新たに z_{ji} , 被説明変量を y_i とすると, y_i と z_{ji} との関係はつぎの線型モデルの形に表現できる。

$$y_i = \beta_0 + \beta_1 z_{1i} + \dots + \beta_j z_{ji} + \dots + \beta_p z_{pi} \dots\dots\dots(8)$$

for $i=1, 2, \dots, N, \quad N=(N_1-2L_1)(N_2-2L_2)$

$$j=1, 2, \dots, P, \quad P=n(2L_1+1)(2L_2+1)$$

ここで, 回帰係数 β_i は残差平方和を最小とする最小二乗推定値として次の方程式を解いて推定することができる。

$$\left. \begin{aligned} S_{11}b_1 + S_{12}b_2 + \dots + S_{1p}b_p &= S_{y1} \\ S_{12}b_1 + S_{22}b_2 + \dots + S_{2p}b_p &= S_{y2} \\ \vdots \\ S_{1p}b_1 + S_{2p}b_2 + \dots + S_{pp}b_p &= S_{yp} \end{aligned} \right\} \dots\dots\dots(9)$$

$$b_0 = \bar{y} - b_1 \bar{z}_1 - b_2 \bar{z}_2 - \dots - b_p \bar{z}_p \dots\dots\dots(10)$$

$$\text{for } S_{ij} = \sum_{k=1}^N (z_{ik} - \bar{z}_i)(z_{jk} - \bar{z}_j), \quad i, j=1, 2, \dots, p$$

$$S_{yi} = \sum_{k=1}^N (y_k - \bar{y})(z_{ik} - \bar{z}_i), \quad i=1, 2, \dots, p$$

$$\bar{z}_i = \left(\sum_{k=1}^N z_{ik} \right) / N, \quad \bar{y} = \left(\sum_{k=1}^N y_k \right) / N$$

このとき, 共分散 S_{ij} , S_{yi} の値は $(N_1-2L_1)(N_2-2L_2)$ 組のペアのデータから計算される値であり, 先に図—1(B) で示したデータ範囲のとり方を M_0 として計算したときの空間共分散関数の値に等しいものとなる。また, その空間共分散関数からなる連立方程式 (9) 式は (5) 式の連立方程式と等価であって, 得られる回帰係数の推定値 b_i を写像 P の逆写像 P^{-1} で逆変換すれば, データ範囲のとり方を M_0 として推定したときの空間影響関数を得ることができる。

以上から, 比較的大規模となる SCA 法の計算プログラムを開発する代わりに, この並び替え写像 P の計算プログラムを準備するだけで, 重回帰分析のパッケージ・プログラムを利用して空間影響関数モデルの推定を行うことが可能となり, 計算プログラム開発に要する労力を軽減することができる。

2.4 FFT 利用による計算の高速化

本項では, SCA 法の基礎となる空間相関関数の計算を迅速かつ正確に行うため, 既存の FFT (fast Fourier transform; 高速フーリエ変換) と呼ばれるフーリエ変換プログラムを活用できることを示す。

通常のフーリエ変換⁽⁵⁾は, 例えば時間の関数としての音量を音の周波数の関数としてとらえなおすというように, 一次元変量をその一次元周波数領域へと移し変える操作であり, 周波数領域 (これと数学的に等価な角速度領域で表すこともある) で簡単に計算できるパワース

ペクトルの逆フーリエ変換が一次元で定義される共分散関数 (相関関数) になるという有名な Wiener-Khintchine の公式があり, これを利用して共分散関数を計算している。このとき, 周波数領域へのフーリエ変換, パワースペクトルからの逆フーリエ変換の計算において FFT が用いられる。その理由は, 巧妙な並び替え操作をすることで演算回数を極端に少なくしているため, 高速であることと, 演算繰返しによる計算誤差が理論的に小さくなるためである。

本論文で対象としているメッシュデータは, 二次元変量であるため, このフーリエ解析のアイデアと FFT の活用のためには, 既存のフーリエ変換の技法が活用できるように, 二次元変量のフーリエ変換という一般化をしなくてはならない。ここで行う提案では, 二次元変量のフーリエ変換という操作を, メッシュ座標の t_1 方向に通常のフーリエ変換を行い, 続いて t_2 方向に通常のフーリエ変換を行うという自然な一般化によって定義することで, 通常のフーリエ解析の考え方がほぼ活用できること, すなわち Wiener-Khintchine の公式の二次元版に相当する公式が成立することを示す。さらに, この公式を用いることで, 原メッシュデータから二度の通常のフーリエ変換を用いて計算した二次元パワースペクトルからの通常の逆フーリエ変換を二度繰返すだけで, 空間相関関数が計算できることを示す。この二度のフーリエ変換と二度の逆フーリエ変換において FFT を活用することを示す。

上記の考え方が, 数学的に可能であることを示すためメッシュデータつまり区間 ($|t_1| \leq T_1/2$; $|t_2| \leq T_2/2$) で定義される二次元変量を $x(t_1, t_2)$ とするとき,

$$\alpha(n_1, n_2) = \frac{1}{T_1 T_2} \int_{-T_1/2}^{T_1/2} \int_{-T_2/2}^{T_2/2} x(t_1, t_2) \\ \cdot \exp[-i 2 \pi (n_1 t_1 / T_1 + n_2 t_2 / T_2)] dt_1 dt_2 \\ \dots\dots\dots(11)$$

なる変量を定義する (本項に限り記号 i は, 変量 X_i を表すサフィクスではなく虚数単位を表す)。このとき, もとの二次元関数を $x(t_1, t_2)$ 自体がこの $\alpha(n_1, n_2)$ を用いて以下のように表されることが簡単に証明できる⁽⁶⁾。

$$x(t_1, t_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \alpha(n_1, n_2) \\ \cdot \exp[i 2 \pi (n_1 t_1 / T_1 + n_2 t_2 / T_2)] \\ \dots\dots\dots(12)$$

さらに, (11) 式は, 角速度 ω を用いるときには,

$$W(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1, t_2) \\ \cdot \exp[-i(\omega_1 t_1 + \omega_2 t_2)] dt_1 dt_2 \\ \dots\dots\dots(13)$$

と表すことができる。この $W(\omega_1, \omega_2)$ を $x(t_1, t_2)$ の二次元フーリエ変換と呼ぶことにする。ここで, イクスボネ

ンシャル内で和の形になっている t_1 と t_2 は積分項内では積の形とみなせ、 t_1 で積分してから t_2 で積分すると考えてよく、この積分計算に相当するのが通常の一次元のフーリエ変換に相当している。したがって、 t_1 に関して通常のフーリエ変換をし続いて t_2 に関して通常のフーリエ変換したものが、ここで (13) 式によって定義された二次元のフーリエ変換ということになる。

このとき (12) 式も角速度表示を用いると

$$x(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\omega_1, \omega_2) \cdot \exp[i(\omega_1 t_1 + \omega_2 t_2)] d\omega_1 d\omega_2 \cdots (14)$$

となる。つまり、 $W(\omega_1, \omega_2)$ の二次元フーリエ変換によって $W(\omega_1, \omega_2)$ を用いて $x(t_1, t_2)$ を求めることができる。このことは通常の一次元フーリエ変換の自然な二次元空間への拡張が可能になったことを示している。

次にパワー・スペクトル $S(\omega_1, \omega_2)$ を

$$S(\omega_1, \omega_2) = \lim_{T_1 \rightarrow \infty} \lim_{T_2 \rightarrow \infty} [(2\pi)^2 W(\omega_1, \omega_2) \cdot W^*(\omega_1, \omega_2) / T_1 T_2] \cdots (15)$$

と定義する。そこで自己相関関数の定義

$$C(\tau_1, \tau_2) = \lim_{T_1 \rightarrow \infty} \lim_{T_2 \rightarrow \infty} \frac{1}{T_1 T_2} \int_{-T_1/2}^{T_1/2} \int_{-T_2/2}^{T_2/2} x(t_1, t_2) \cdot x(t_1 + \tau_1, t_2 + \tau_2) dt_1 dt_2 \cdots (16)$$

から、(14) 式を代入し積分順序を入れ換える。このとき積分区間は $(-\infty, \infty)$ となり、さらに共役関係を利用すると以下の二次元空間における Wiener-Khintchine の公式を導くことができる。

$$C(\tau_1, \tau_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\omega_1, \omega_2) \cdot \exp[i(\omega_1 \tau_1 + \omega_2 \tau_2)] d\omega_1 d\omega_2 \cdots (17)$$

$$S(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C(\tau_1, \tau_2) \cdot \exp[-i(\omega_1 \tau_1 + \omega_2 \tau_2)] d\tau_1 d\tau_2 \cdots (18)$$

つまり、二次元空間における相関関数とパワー・スペクトルは互いにフーリエ変換の関係にあり、一方が知れば他方も求めることができる。

以上から、図-3 に示すように、SCA 法の空間相関関数を計算する過程において、まず、メッシュデータの情報を二次元スペクトル領域へ二次元フーリエ変換し、求められたスペクトルから二次元パワー・スペクトルを計算し、次に、二次元の Wiener-Khintchine の公式 (17) および (18) 式を利用してパワー・スペクトルの二次元逆フーリエ変換を行えば、直ちに、空間相関関数を求めることができる。

この際、二次元フーリエ変換の計算において、メッシュデータの範囲 ($N_1 \times N_2$) が、 $N_1 = 2^h$ 、 $N_2 = 2^m$ (h, m は正の整数) となっていれば、FFT のパッケージ・プログラムを利用して理論的にも計算誤差を押さえた計算が可

図-3 二次元空間におけるフーリエ変換と Wiener-Khintchine の公式

能となり、直接計算する場合よりも演算時間の飛躍的な短縮を図ることが可能となる。

3. 空間相関関数の信頼性と視覚化

3.1 相関関数の確率分布

既に報告したように^[2]、SCA 法における空間相関関数を用いれば、都市活動要素の空間的影響関係を計量的に解析することが可能となる。しかしながら、1 回のコイン投げの実験で表が出たからといって裏が出る確率は 0 と結論することは危険なように、具体的なデータから計算される空間相関関数の値を、普遍的な数字として扱うことには危険性がある。限られた地域のメッシュデータを扱う以上、SCA 法を利用する際にこの危険性を避ける方法を確認しておく必要がある。コイン投げの実験では、表の出る確率は、実験回数と表の出た頻度から決まるある範囲におさまっていると判断してよいというような統計的推論の方法があるが、本研究でも、この統計学的アプローチを用いて、空間相関関数の値の信頼区間を求める理論を構築しておきたい。

相関関数 ρ が確率的に変動しており、これを母集団とし、ここから有限の標本が得られたとし、その標本相関関数 r の分布から、元の相関関数 ρ についての検定や推定を行うことを検討する。しかし、この場合、確率分布についての情報が必要とされ、特に元の相関関数 ρ が 0 とみなせない場合は確率分布は大変複雑となることが知られている。そこで、次に示すフィッシャーの Z 変換と呼ばれる変換^[7]を施すと、もとの確率変量の分布の情報がわからない場合でも、 Z 変換した量は、正規分布に従うという性質があり、正規分布に対して統計学で構築されている検定や推定の理論を活用することが、以下に示すように可能になってくる。

標本相関関数 r の Z 変換値

$$z = \frac{1}{2} \log_e [(1+r)/(1-r)] \cdots (19)$$

は、ほぼ平均 ζ 、分散 σ_z の正規分布に従うという性質が知られており、また、これらの量 ζ, σ_z も次のように、相関関数 ρ を用いて

$$\zeta = \frac{1}{2} \log_e [(1+\rho)/(1-\rho)] \cdots (20)$$

$$\sigma_z = 1/\sqrt{n-3} \quad (n: \text{標本数}) \cdots \cdots (21)$$

と表せることが知られている。

したがって、標本数 n が十分大きいとき (例えば $n > 40$)、この Z 変換を用いた正規分布近似を利用して以下の項に示すように相関関数の値の検定や推定を行うことが可能となる。

3.2 空間相関関数の区間推定論

元の変量の確率分布がどのような場合でも、上記の Z 変換した値は正規分布に近似可能なことから、推定値の区間推定に利用することができる。例えば、95 % で信頼できる (誤った判断すなわち以下で示される範囲の外になってしまう率 (これを危険率と呼ぶこともある) が 5 % ということであり、この 95 % の方を信頼係数と呼ぶこともある) 母集団相関関数 ρ の Z 変換値 ζ の範囲を知るには、95 % の範囲が標準正規分布では $(-1.96, 1.96)$ の範囲になることから、

$$\zeta_1 = z - 1.96 \sigma_z \leq \zeta \leq z + 1.96 \sigma_z = \zeta_2 \cdots \cdots (22)$$

となる。 z , σ_z は (19), (21) 式によって元のデータから計算できる。さらに、 ζ の逆 Z 変換、

$$\rho = [\exp(2\zeta) - 1] / [\exp(2\zeta) + 1] \cdots \cdots (23)$$

を ζ_1 , ζ_2 に施し、その値を ρ_1 , ρ_2 とすれば、

$$\rho_1 \leq \rho \leq \rho_2 \cdots \cdots (24)$$

となる。すなわち、母集団相関関数 ρ の 95 % 信頼区間は ρ_1 から ρ_2 の範囲であることが分かる。この範囲が広ければ ρ の推定値は、この範囲の中でふらついているということになり、1 回だけのコイン投げの例のように推定値をそのまま信じることは危険である。このようなことは、データ数の少ないときに生じやすく、メッシュデータを得る対象領域が狭い場合には、(24) 式の範囲を常に考慮して判断すべきである。

具体的な数値例として、大阪府の土地利用メッシュデータ⁽¹⁸⁾における商業・業務地と 1) 一般市街地、2) 社寺敷地・公園庭園の組み合わせから空間相関関数を計算し、上記の理論を用いて、信頼係数 0.95 のもとの空間相関関数の値の区間推定を行い、結果を示したものが図—4.1, 4.2 (前報⁽²⁾) の図—6.1, 6.3 に対応している) で、空間相関関数の値を空間ラグ (k_1, k_2) 平面に対し垂直の方向にとり、同時にその値の信頼区間も示している。図のような視覚的表現によれば、都市活動要素の空間的影響関係を容易にイメージでき、前報⁽²⁾で示した商業業務と一般市街地の排斥性、商業業務と社寺庭園の共存性が、推定値の有効範囲も同時に見ながら判断することが可能となる。

3.3 空間相関関数の値の差の有意性判定

図—4.1, 4.2 に示したように、都市活動要素の空間的影響の様子が空間ラグ (k_1, k_2) 上の空間相関関数の値の変化によって見る事ができる。しかし、空間相関関数の値は、上述のように統計学にはある範囲にある量と

して認識すべきであり、変化しているように見える空間相関関数の値が統計的に有意な変化であるか否かを検証する必要がある。そのため、はじめに空間相関関数の値の差の有意性検定のための基礎的理論について整理しておく。

ある変量 X_j の要素 z と他の変量 X_i の 2 つの要素 x , y との間の相関係数 r_{xz} , r_{yz} の差の有意性を検定する場合には、つぎに示す方法によらなくてはならない⁽¹⁹⁾。まず、標本から得られる相関係数から、

$$t = (r_{xz} - r_{yz}) / \sqrt{(n-3)(1+r_{xy})/2(1-r_{xy}^2-r_{xz}^2-r_{yz}^2+2r_{xy}r_{xz}r_{yz})} \cdots \cdots (25)$$

なる統計量 t を求めると、この統計量 t は母集団相関係数が同じという条件 ($\rho_{xz} = \rho_{yz}$) のもとで、自由度が $n-3$ (n は標本数) の t 分布をすることが知られている。この理論を利用すれば、2 つの相関係数の差の有意性を検定することができる。つまり、ある有意水準での限界値を t_a とするとき、得られる t 値が

$$-t_a \leq t \leq t_a \cdots \cdots (26)$$

となれば $\rho_{xz} = \rho_{yz}$ という帰無仮説は棄却されず、2 つの相関係数の値に有意な差を認めることはできないと判定することができる。

3.4 空間相関関数の構造化

上記の理論において、空間相関関数の値を

$$\left. \begin{aligned} r_{xz} &= R_{ij}(s_1, s_2) \\ r_{yz} &= R_{ij}(t_1, t_2) \\ r_{xy} &= R_{ii}(s_1 - t_1, s_2 - t_2) \end{aligned} \right\} \cdots \cdots (27)$$

のように対応づければ、空間相関関数の値の差の有意性を検定することができる。ただし、前述した $M(k_1, k_2)$ の範囲からデータをとるときには標本数 n 、つまり空間共分散関数や空間相関関数を計算する際に用いるペアとなるデータ数 n は空間ラグによって異なる。そこで、より厳しい検定を行うため r_{xz} , r_{yz} , r_{xy} のそれぞれを計算する際のデータ数の中で最小のものを n として用いる。

ここで、統計量 t の値が (26) 式の範囲に入っていれば、その有意水準において $R_{ij}(s_1, s_2)$ と $R_{ij}(t_1, t_2)$ の母集団相関関数の値には有意な差を認めず、両者は同一のグループであると判断し、この操作を各空間ラグの組み合わせについて繰り返し行うことで、空間相関関数の値のグルーピングを行うことができる。また、各グループの代表値は、各グループに属する空間相関関数の値の確率分布を先の Z 変換によって正規分布に近似させ、そこで相加平均を求めた後、その値を逆 Z 変換する方法によって求めることができる。

空間相関関数の値を、値の差の有意性に注目してグルーピングし、各グループの代表値によって空間的な影響関係を把握しようとするこの方法によれば、後に具体

図—4.1 「商業・業務地」の面積と「一般市街地」の面積との相互空間相関関数と95%信頼区間

図—4.2 「商業・業務地」の面積と「社寺敷地・庭園公園」の面積との相互空間相関関数と95%信頼区間

(グルーピング前の空間相関関数)

(グルーピング後の空間相関関数)

図—5.1 「商業・業務地」の面積と「一般市街地」の面積との相互空間相関関数の構造化

(グルーピング前の空間相関関数)

(グルーピング後の空間相関関数)

図—5.2 「商業・業務地」の面積と「社寺敷地・庭園公園」の面積との相互空間相関関数の構造化

例で説明する図—5.1, 5.2のように内在する統計的に有意な空間相関関数の構造を明示することが可能となる。

3.5 空間相関関数の構造の視覚化

上述の理論を用いて、3.2で用いた同じデータに対してグルーピングを行い、視覚化したものが図—5.1および図—5.2である。空間相関関数の値の差の有意性判定では、有意水準が0.05での両側検定を行っている。 t 分布の限界値 $t_{0.975}$ は自由度 $(n-3)$ によって異なるが、自由度が100のとき $t_{0.975}=1.98$ であり、自由度が無限大で $t_{0.975}=1.96$ であるので、標本数 n が十分大きい本分析では $t_{0.975}=1.97$ としている。両図では、グルーピ

ングする前の様子と併せてグルーピングされた空間相関関数の様子を視覚的に表現している。図のように空間相関関数の構造の視覚化を図ることによって、空間相関関数の変化と同時に、前述の商業業務と一般市街地の排斥性の顕著な近傍領域（図—5.1中のハッチングが薄い領域）、商業業務と社寺庭園の共存性の顕著な近傍領域（図—5.2中の濃いハッチング領域）が、視覚的に把握可能となる。

なお、以上までの計算および視覚情報のプレゼンテーションのため、本研究では、主要な計算を大型計算機で行い¹⁰⁾（駆動は端末から）、結果の出力の内視覚情報に

ついて、白黒の線画の出力は大型計算機に直結するレーザープリンター^{注11)}および端末マイクロコンピュータを介して研究室のXYプロッターに、カラー画像は、端末マイクロコンピュータおよび画像処理装置^{注12)}に出力させている。

4. まとめ

SCA法の計算方法は、計算プログラムとしては比較的大規模なものとなり自己開発するには大きな労力を要するため、その理論体系を再検討することで計算過程の簡略化や計算時間の短縮化を図る方法について検討した。また、実用上の要請として、空間的な影響関係が容易に把握でき、推定値の信頼できる範囲をも同時に判断できる必要があることから、空間相関関数の値の信頼区間を求める方法、値が等しい(異なる)とみなせるかどうかを判断するための検定論という推定値の評価を行う方法を構築した。また、この方法をもとに構造化を行う方法について検討し、さらに、空間的影響関係を視覚的に表示する方法を求めた。つまり、

- 1) 空間影響関数モデルの推定の際には、データに写像 P をほどこす並び替えプログラムを用意することで、直ちにパッケージ・プログラムの利用が可能となり計算労力の軽減を図ることができる。
- 2) メッシュデータが $N_1 \times N_2 = 2^h \times 2^m$ (h, m は正の整数) の形をしているとき、空間相関関数を直接計算する過程に代えて、二次元FFTを利用すれば相対的に演算時間は短縮され計算時間の短縮を図ることができる。
- 3) 空間相関関数の推定値を評価するための信頼区間、検定方法を確立し、グルーピングによる構造化を行う方法について検討し、都市活動要素の空間的影響関係やその推定値の有効範囲を視覚的に判断する方法を提案した。

注

- 1) 文献1) 参照
- 2) 文献2) 参照
- 3) 文献3) 参照
- 4) (a) と (b) の計算方法の違いによって活用データ数に差が最も大きくなるのは、空間ラグが (0, 0) すなわち同一場所での都市活動間の相関であり、後述の例で用いるメッシュデータの場合には、(a) の方法ではデータの数 が 1080 であり、(b) の方法では、714 となっている。このとき、計算例に用いる変量について言えば、都市中央部と周辺部では、相互の影響の仕方が大きく異なると考えられる「商業・業務地」と「一般市街地」との空間ラグが (0, 0) での相関関数の値は、(a) では $C_a = -0.239$ であり (b) では $C_b = -0.323$ とかなり異なる。一方、「商業・業務地」と「社寺敷地・庭園公園」では、空間ラグ (0, 0) での相関関数は、(a) では $C_a = 0.104$ であり (b)

では $C_b = 0.090$ となりほぼ一致した値となる。このように、(a) と (b) との方法による推定値の差は、都市活動を表す変量の特徴にも依存していることに注意する必要がある。いずれの方法においても、後述の推定値の信頼区間を調べる方法によって、その区間が広がっていれば、推定値そのものが相当ふれていると判断し、対象地域の選び方も含めて (a) と (b) の方法を再検討する必要がある。筆者らは、「商業・業務地」と「一般市街地」の例のような場合には、対象地域を細区分し直し、小領域ごとに (a) の方法で計算することが望ましいと考えている。

- 5) たとえば、文献5) にフーリエ変換およびFFTに関する標準的なフーリエ解析の理論と応用が述べられている。
- 6) 通常のフーリエ解析の理論から

$$x(t_1, t_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \beta(t_1, n_2) \exp[i 2 \pi n_2 t_2 / T_2] \cdots \cdots (i)$$

$$\beta(t_1, n_2) = \sum_{n_1=-\infty}^{\infty} \alpha(n_1, n_2) \exp[i 2 \pi n_1 t_1 / T_1] \cdots \cdots (ii)$$

と書ける。そこで

$$\alpha'(n_1, n_2) = \frac{1}{T_1 T_2} \int_{-T_1/2}^{T_1/2} \int_{-T_2/2}^{T_2/2} x(t_1, t_2) \cdot \exp[-i 2 \pi (n_1 t_1 / T_1 + n_2 t_2 / T_2)] dt_1 dt_2 \cdots \cdots (iii)$$

なる n_1, n_2 の関数 α' を想定する。そこで

$$\begin{aligned} & \frac{1}{T_2} \int_{-T_2/2}^{T_2/2} x(t_1, t_2) \exp[-i 2 \pi n_2 t_2 / T_2] dt_2 \\ &= \frac{1}{T_2} \int_{-T_2/2}^{T_2/2} \sum_{n_2=-\infty}^{\infty} \beta(t_1, n_2) \exp[i 2 \pi (n_2 - m_2) t_2 / T_2] dt_2 \\ &= \sum_{n_2=-\infty}^{\infty} \beta(t_1, n_2) \frac{1}{T_2} \int_{-T_2/2}^{T_2/2} \exp[i 2 \pi (n_2 - m_2) t_2 / T_2] dt_2 \\ &= \beta(t_1, m_2) \\ &= \sum_{n_1=-\infty}^{\infty} \alpha(n_1, m_2) \exp[i 2 \pi n_1 t_1 / T_1] \cdots \cdots (iv) \end{aligned}$$

であることを利用すると、(iii) 式より

$$\begin{aligned} \alpha'(m_1, m_2) &= \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} \left[\sum_{n_1=-\infty}^{\infty} \alpha(n_1, m_2) \exp[i 2 \pi n_1 t_1 / T_1] \right] \\ &\quad \cdot \exp[-i 2 \pi m_1 t_1 / T_1] dt_1 \\ &= \sum_{n_1=-\infty}^{\infty} \alpha(n_1, m_2) \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} \\ &\quad \cdot \exp[i 2 \pi (n_1 - m_1) t_1 / T_1] dt_1 \\ &= \alpha(m_1, m_2) \cdots \cdots (v) \end{aligned}$$

以上のことから (12) 式における係数は (11) 式より計算可能である。

- 7) たとえば、文献6) の112頁にフィッシャーのZ変換と変換値が正規分布するという性質が述べられている。
- 8) 大阪府の土地利用の面積比率のデータであり、メッシュサイズは東西方向約574m、南北方向約463m、メッシュ数は1080(27×40)個である。文献2)、文献3)で用いたデータと全く同じものを使用している。
- 9) 変量 x と変量 y に関しては特定の分布形を仮定しないが、変量 z に関しては『変量 z は変量 x の固定した値に対してすべて正規分布をし、その分散は x のいずれの値についても等しいものとし、また y の固定した値に対しても同様である』という仮定が必要であり、 x, y の値をそこに得られた標本値の組に限定し、 z だけを確率変数とみなして行われる。
- 10) 大型計算機としては、現在まで国内メーカー2機種、海

外メーカー1機種上で作動させているが、プログラムの移植に伴うトラブルは全くなく、計算時間も今回報告した計算例でデータ入力から空間相関関数の計算、空間影響関数モデルの推定、各変量のモデルによる推定まで行ってCPUで5.6秒前後であり問題なく、メモリーサイズ制約にかかる問題は生じていない。また、本研究では既存パッケージを利用する際、SAS（文献7）参照）を利用している。計算時間、メモリーサイズ制約にかかることはないが、出力用のワークエリアを通常よりも大きく指定する必要がある。

- 11) 線画的出力に関しては、現在の段階では機種によって指定の仕方が異なる場合があるので、その都度機種のマニュアルにそってプログラムを若干変更させる必要がある場合がある。
- 12) カラー画像出力に関しても機種によって指定の仕方が異なるため、端末を介して間接的に表示させているが、現在普及しつつあるワークステーションレベルの機種が利用できれば、全体の計算からカラー画像出力まで一貫的にかつインタラクティブに行うことが可能となる。

文献リスト

- 1) 青木義次：メッシュデータ解析の一方法としての空間相

関分析法の提案—その1 メッシュデータ解析の問題点と空間相関分析法の理論，日本建築学会論文報告集，第364号，pp.94～101，1986年6月

- 2) 青木義次：メッシュデータ解析の一方法としての空間相関分析法の提案—その2 土地利用の連担性・共存性・排斥性の計量化への応用，日本建築学会論文報告集，第368号，p.119～125，1986年10月
- 3) 青木義次：メッシュデータ解析の一方法としての空間相関分析法の提案—その3 空間影響関数モデルの有効性と問題点，日本建築学会論文報告集，第377号，p.29～35，1987年7月
- 4) 青木義次，大佛俊泰：空間相関分析法の計算方法，日本建築学会，第10回電子計算機利用シンポジウム論文集，pp.79～84，1988年3月
- 5) 日野幹雄：スペクトル解析，朝倉書店，1977
- 6) 芝 祐順，渡部 洋：統計的方法Ⅱ，新曜社，1976
- 7) 雄山真弓ほか：SASによるデータ解析（上・下），丸善，1984

（1990年2月10日原稿受理，1990年7月31日採用決定）