

SFI 規準による空間分布モデルのためのパラメータ推定法*

A Parameter Estimation Method based on the Spatial Fit Indicator for Spatial Distribution Models*

宮本 和明**, 内田 敬***, 明石 和之****

By Kazuaki MIYAMOTO, Takashi UCHIDA and Kazuyuki AKASHI

1. はじめに

分布交通量, 土地利用モデルなどの空間分布量を対象とした都市モデルにおいては, 対象地域を多数のゾーンに分割し, 集計されたデータをもとに個々のゾーンの特性量の予測分析が一般に行われている。

統計モデルの推定法として多く用いられている通常最小 2 乗法(OLS)においては, 各ゾーンの値は独立, 誤差は均等分布と仮定されているが, 空間データに OLS を適用した場合, 観測値とモデルによる予測値の残差が一定地域に偏ってしまい空間分布として観測値をうまく再現できない場合がある。本来, 人口分布などの空間分布においてはある地域に集中するような独自の空間パターンが存在し, ゾーンは互いに従属関係にあると考えられ, OLS における独立性の仮定を満たさないと考えられる。

誤差が空間的に連担する場合, 通常最小 2 乗法によると推定量の分散がバイアスを持ち, 誤った統計的推測を導く。これを誤差の空間的自己相関問題と呼び, 対処するためのパラメータ推定法・検定法が多く提案されている^{1),2)}。また, 分布の不均一性の診断指標の提案, 拡張も活発に行われている^{3),4),5),6)}。

一方, オペレーショナルな観点からは空間データを取り扱う予測モデルの予測性能, すなわち観測値と予測値の適合度が重視される。線形モデル

の OLS 推定では相関係数 (あるいは決定係数) が適合度として通常用いられるが, 空間的自己相関が存在する場合には, この指標では不適切であることが指摘されている⁷⁾。そこで, 適合度指標の多様性を許す柔軟なパラメータ推定方法として, 線形計画法や目標計画法を応用する手法も提案されている⁸⁾。しかし, これらはいずれも, ゾーンは独立との前提で観測値と予測値を各ゾーン 1 対 1 対応として適合度評価するものである。

しかし, 都市モデルのようなゾーン単位の集計データを取り扱う統計モデルにおいては, 分析データそれ自身がゾーンの境界の設定に依存し, 観測値自体が揺らぎを持つ。そのため, 各ゾーン値の適合を評価するだけでは不十分で周辺ゾーンを考慮した適合度評価が必要である。

上述の問題意識から提案されているのが空間分布適合度指標(SFI)である⁹⁾。本研究では, さらに空間分布モデルのパラメータ推定の一方法として, ゾーンの空間的位置関係を考慮するために SFI に基づいた推定法を提案する。本研究の SFI 規準パラメータ推定法は, LP (線形計画法) によりパラメータ推定問題を定式化することによりモデルの適合度をモデルの目的に応じて柔軟に考えていく点で O'Kelly et al.⁸⁾と基本的な視点を同じくする。適合度規準として SFI を用いており, SFI の算出自体が LP 問題であるため, 適合度算出とパラメータ推定の 2 つの LP 問題となるところを, 1 問題として解くところに特徴がある。

2. 空間分布適合度指標

2.1 空間分布における適合度の考え方

2 次元空間分布モデルにおいては, ゾーン境界の設定という手続きのためにゾーンの観測値自体が揺らぎを持つ。従って, 観測値と予測値の対応関係を 1 対 1 では考えず, 予測値の残差はその生じているゾーンの周辺ゾーンで吸収可能であれば

* キーワーズ: 空間分布, SFI, パラメータ推定

** フェロー, 工博, 東北大学教授, 東北アジア研究センター
(〒980-77 仙台市青葉区荒巻字青葉
Tel: 022-217-7475, Fax: 022-217-7477
E-mail: miyamoto@plan.civil.tohoku.ac.jp)

*** 正員, 工博, 東北大学助教授, 大学院工学研究科土木工学専攻

(Tel: 022-217-7476, Fax: 022-217-7477
E-mail: uchida@plan.civil.tohoku.ac.jp)

**** 学生員, 東北大学大学院情報科学研究科
(Tel: 022-217-7479, Fax: 022-217-7477
E-mail: akashi@rs.civil.tohoku.ac.jp)

構わないとする。そこで、残差の大きさのみならず、その空間分布に着目する。例えば、正負の残差が隣接していれば、それは観測値自体の境界依存問題が影響している可能性があるが、同符号の残差が空間的に連担している場合にはモデルが不適切と判断すべきである。従って、適合度評価においては、残差の大きさもさることながら、正、負の残差分布の関連度も考慮しなければならない。SFI では、予測値分布において残差の正負を相補うような移送を考え、移送距離をもって分布の関連度を計量化する。

2.2 空間分布適合度指標(SFI)⁹⁾

2 つの空間分布、例えば観測値分布とモデルによって推定された予測値分布において、各ゾーン値の総量が等しい場合には、予測値分布から予測値と観測値の差、すなわち「残差」を他のゾーンに移送することによって観測値分布を再現することができる。このとき、残差の移送距離を考慮し、残差の絶対値と移送距離の積の総和である「総移送費用」を指標(SFI)として空間分布の適合度を評価するものである。

SFI は以下のように算出される。観測値分布と予測値分布の要素の総量が等しいとし（前提条件）、過大予測ゾーンを残差の供給地、過小予測ゾーンを残差の需要地と見なすと、残差の移送は、残差 T を操作変数、総移送費用を目的関数 S とする、線形計画の輸送問題として定式化できる。SFI は目的関数 S の最小値として求められる。

$$S = \sum_{i,j} d_{ij} \cdot T_{ij} \rightarrow \min. \tag{1}$$

s.t.

$$\sum_j T_{ij} \leq a_i, \quad \forall i \tag{2}$$

$$\sum_i T_{ij} \geq b_j, \quad \forall j \tag{3}$$

$$T_{ij} \geq 0, \quad \forall i,j \tag{4}$$

ここに、

- T_{ij} : ゾーン i から j への移送量
- d_{ij} : ゾーン i, j 間の移送距離
- a_i : 過大予測ゾーン i の過大量
- b_j : 過小予測ゾーン j の過小量

2.3 数値計算例

3×3 の簡単な分布(Black⁷⁾)に対する 2 つの予測値分布（残差が凝集している分布、凝集していない分布）を用いて残差分布と SFI 値との対応関係を見る。表.1 には計算に用いた分布を、表.2 には代表的に用いられる適合度指標値、図.1、図.2 に、2 つの予測 A, B それぞれの残差分布を示す。

表.1 数値計算例

観測値分布					
5	10	20			
40	75	100			
95	115	80			
予測値分布 A			予測値分布 B		
15	20	0	15	20	0
60	85	70	60	85	70
145	95	50	65	95	130

表.2 適合度評価

	分布 A	分布 B
r	0.8126	0.7936
RMSE	25.386	25.386
SFI	320.000	228.284

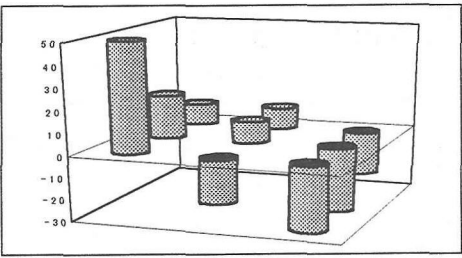


図.1 残差分布 A

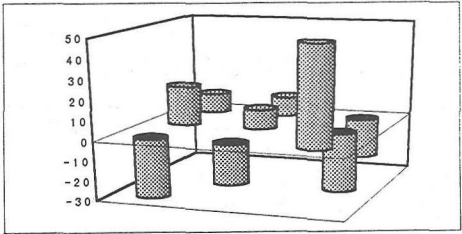


図.2 残差分布 B

表.2 によると、RMSE(平方根平均 2 乗誤差)は両方同じだが、各ゾーンを独立に評価対象とする r (相関係数) は残差凝集分布 A の評価が高い。

方, SFI 値は小さいほうが望ましいことから残差凝集分布 A が分布 B に比べ 30%程度劣っていることとなり, r と逆の評価となっている。

SFI は残差分布がランダムであれば, 残差の絶対値に比例的な, RMSE と同様の特性を示す。しかし, 残差分布がランダムでない場合には, 各ゾーンの残差の絶対値及び同符号の残差の連坦度が考慮されていることが判る。

3. SFI 規準/パラメータ推定法

3.1 SFI 規準推定の意義

一般に統計モデルによる予測・分析過程では, 1)探索的→2)確認的→3)適応的, の 3 局面が考えられる¹⁰⁾。SFI 規準推定法は, モデル推定の 3)適応段階で用いることを想定する。従って, 残差分析等によって現象説明に十分な最適なモデル候補が定まってい, それらの具体的なパラメータ推定と, 最終的なモデル選択が課題となる段階で用いる。このとき, パラメータ推定値の性質よりもその実際の適用上の機能性, すなわち, 予測値分布の適合度を重視する。予測の適合度評価は種々の指標が用いられ得るが, その一つとして, サンプル点(ゾーン)相互の空間配置を考慮した SFI を適合度評価として用いるものである。

3.2 定式化

ここでは, 予測モデルが一般的な線形回帰式で表される場合について考える。

SFI 規準パラメータ推定法は, SFI の定式化(1)~(4)において, 予測モデルのパラメータを線形計画法の操作変数として新たに付加して, 各ゾーンの観測値と予測値の差を最も効率よく移動させた場合の「総移送費用」, すなわち SFI を最小とするようにパラメータを求めるものである。

$$S = \sum_{i,j} d_{ij} \cdot T_{ij} \rightarrow \min. \quad (5)$$

s.t.

$$\sum_j T_{ij} \leq Y_i, \quad \forall i \quad (6)$$

$$\sum_i T_{ij} - \sum_i T_{ji} + Y_i \geq A_j, \quad \forall j \quad (7)$$

$$\sum_i Y_i = \sum_i A_i \quad (8)$$

$$T_{ij} \geq 0, \quad \forall i, j \quad (9)$$

$$Y = X\beta \quad (10)$$

ここに,

A_i : ゾーン i の観測値

Y : $(n \times 1)$ 内生変数

X : $(n \times m)$ 外生変数

β : パラメータ

パラメータ推定問題では, 予測値と観測値の大小関係は未定であるため線形計画問題の制約条件式が SFI とは異なる。すなわち, パラメータ推定問題の需要条件式(7)が SFI の制約式(2), (3)に対応する。予測値分布は観測値分布の配分問題と考えられるため, (8)式において総量一致としている。(6)式は, 計算の便宜上導入される凡長な条件式である。

式(10)では単純な回帰モデルとしているが, これをより複雑な線形都市モデル, 例えば空間相互作用を考慮したモデル(11)や, 空間的自己回帰モデル(12)に拡張することができる。また, 適用例に示すように, 重力モデル (エントロピー最大化モデル) も変数変換により線形化できることから, 本推定法は都市モデル全般に広く適用可能である。

$$Y = X\beta + WX\rho \quad (11)$$

$$Y = \zeta WY + X\beta \quad (12)$$

4. 適用例

平成 4 年度仙台都市圏パーソントリップ調査データを用いて, 仙台市中心ゾーン (図 3: 黒地域) を終点ゾーンとする, 内生変数間の相互作用, 空間相互作用を考慮しない単純な OD 分布交通量モデルのパラメータ推定を重回帰分析と SFI 規準パラメータ推定方法によって行い比較検討した。

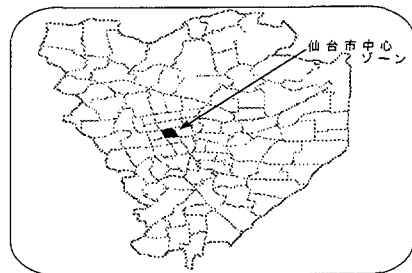


図.3 対象地域

SFI 規準パラメータ推定法では予測モデルは線形形式を考えているため、予測モデル(13)の両辺の対数をとって線形化する。推定結果を表3に、中心ゾーンからの距離に沿っての分布形を図4に示す。

OD 分布交通量モデル：

$$Y_i = \alpha P_i^\beta \exp(-\gamma D_i) \quad (13)$$

ここに、

Y_i ：ゾーン*i*から中心ゾーンへの交通量

P_i ：ゾーン*i*における夜間人口

D_i ：ゾーン*i*から中心ゾーンへの距離

α, β, γ ：パラメータ

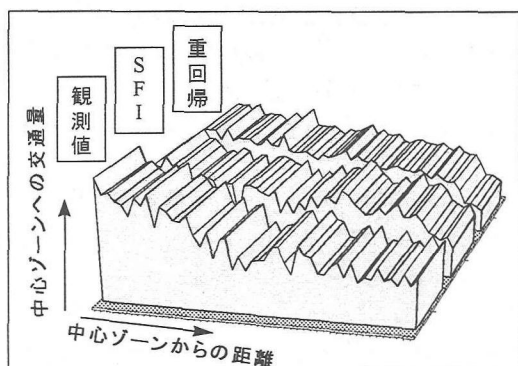


図4 OD 分布型

表3 パラメータ推定結果

パラメータ	説明変数	重回帰	SFI規準
α	定数項	7.7312	0.7602
β	夜間人口	-0.2946	0.6548
γ	距離	0.0491	0.0711
	相関係数	0.7374	0.7191
	SFI	462.97	429.73

表3によると SFI 規準による予測値分布は、重回帰分析による予測値分布に比べ SFI が 10%程度向上しており、ゾーンの位置が考慮された分布が再現されている。また、符号条件も合致している。

図4では、重回帰分析による分布は平均的な値となり観測値分布を再現しきれていないが、SFI 規準による分布形は観測値分布の凹凸を再現しておりゾーンの位置が考慮されていることが判る。

5. まとめ

本稿では、空間分布モデルに対する SFI 値の意味づけを行い、SFI と残差分布の対応関係を示した。また、SFI に基づき空間分布特性を考慮した SFI 規

準パラメータ推定法を提案し適用例を示した。

適用例により内生変数間相互作用、空間相互作用を考慮しない単純なモデルにおいても、SFI 規準パラメータ推定によれば、重回帰分析による分布よりも SFI 値が向上し、空間分布を適切に再現することが確認できた。

今後は、具体的な適用を進めて、空間相互作用を考慮したモデルに SFI 規準推定を適用した場合のパフォーマンスを明らかにする予定である。

<参考文献>

- 1) Florax, R and H. Folmer : Specification and estimation of spatial linear regression models, Regional Science and Urban Economics 22 , pp.405-432, 1992.
- 2) Anselin, L and S. Rey: Properties of tests for spatial dependence in linear regression models, Geographical Analysis, Vol.23, No.2, pp.112-131, 1991.
- 3) Nass, D and D. Garfinkle: Localized autocorrelation diagnostic statistic (LADS) for spatial models — conceptualization, utilization, and computation, Regional Science and Urban Economics, Vol.22, pp333-346, 1992.
- 4) Nass, D and Y. Moon: Localized autocorrelation diagnostic statistic (LADS) for social models — time-series, network and geographic data sets, Sociological Methods and Research, Vol.25, No.2, pp.223-285, 1995.
- 5) Ord, J.K. and A. Getis: Local spatial autocorrelation statistics: Distributional issues and an application, Geographical Analysis, Vol.27, No.4, pp.286-306, 1995.
- 6) Anselin, L. et al.: Simple diagnostics tests for spatial dependence, Regional Science and Urban Economics, Vol.26, pp77-104, 1996.
- 7) W. R. Black : A note on the use of correlation coefficient for assessing goodness-of-fit in spatial interaction models, Transportation, Vol.18, pp. 199-206, 1991.
- 8) M. E. O'Kelly et al.: New estimates of gravitational attraction by linear programming, Geographical Analysis, Vol.27, No.4, pp271-285, 1995.
- 9) 宮本和明, 三浦良平：空間分適合度指標 (SFI) の提案, 土木計画学研究・論文集, No.10, pp.135-142, 1992.
- 10) 竹内啓編：統計学事典, p.459, 東洋経済新報社, 1995.