

誤差項に空間的自己相関が存在する回帰モデルのパラメータ推定手法に関する考察*

*A Study on the Statistical Inference Methods for Regression Models in the Presence of Spatial Autocorrelation**

堤盛人**・清水英範***・福本潤也****・井出裕史*****

Morito TSUTSUMI**・Eihan SHIMIZU***・Jun-ya FUKUMOTO****・Hiroshi IDE*****

1. はじめに

社会資本整備が地域へ及ぼす諸影響を客観的かつ定量的に分析するために、地価モデルや交通・立地モデルといった空間計量モデル（地域モデル）が用いられる。空間計量モデルの適用に際しては、空間的依存性・空間的異質性・可変地区単位問題といった空間概念に起因する統計学的問題がモデルの推定に影響を及ぼし、これに対する適切な処置を欠くとモデルを用いた政策分析等の信頼性が低下してしまうことが知られている¹⁾⁴⁾。

本研究ではこれらの問題のうち、特にクロスセクションモデルのパラメータ推定における空間的依存性の問題に着目し、これによって空間的自己相関と呼ばれる誤差項の系列相関が引き起こされる場合の統計学的対処法について空間計量経済学や計量地理学等の分野で研究されている手法を整理して示す。

これらの手法は、一概にどれが一番優れているとは言えないため、手法の選択は分析者の判断に委ねられることとなる。そのため、適用する手法によって分析結果が大きく異なるのか否かということは、政策の影響分析を行う際の大きな関心事となる。これまで、いくつかの手法を比較分析した研究はあるが、多くは一般的特性を把握することを目的としてシミュレーションデータを用いたものであり⁵⁾⁶⁾⁷⁾、また、具体的なデータを用いた実証分析においても、地区の犯罪率⁸⁾や発病率⁹⁾といった現象を取り扱っており、社会資本整備に密接に関連した事例はあ

まり多くない。そのため、プロジェクトの影響分析においてこれらの統計学的問題を軽視することが、モデルの推定やこれを用いた分析結果にどの程度影響を与えるかということに関しては、研究者の間で必ずしも認識が一致していないと思われる。

そこで本研究では、社会基盤整備プロジェクトの影響分析において特に重要な要因となる地価の分析を例として、誤差項に空間的自己相関が生じる回帰モデルに対していくつかの統計学的対処法を適用し、それらが分析結果に及ぼす影響について、モデルのパラメータ推定値やモデルによる予測値並びにそれらの分散等の観点から実証比較を行うこととする。

2. 空間概念に起因する統計学的問題

空間概念に起因する統計学的問題として、空間的依存性、空間的異質性、可変地区単位問題があげられる。これらについて簡単に説明する。

(1) 空間的依存性

クロス・セクションデータに対して回帰分析などを行う場合、データの測定問題やモデルから抜け落ちたデータ間の外部効果や波及効果のためにデータサンプルの誤差項同士に相関が残る、すなわち空間的自己相関 Spatial Autocorrelation が存在することがある。また、空間相互作用や空間拡散過程といった空間過程 Spatial Process をモデルが十分に説明しきれない場合、近隣ゾーンの被説明変数間にある関数関係が存在することがある。Anselin ら¹⁾²⁾は、それらを総称して空間的依存性 Spatial Dependence と呼んでいる。空間的依存性が存在する場合、これを無視して通常最小二乗法 Ordinary Least Squares Method : OLS による推定を行うと、推定量が偏りを持ったり、あるいは OLS 推定量は不偏でも有

*キーワード：整備効果計測法、調査論、地価分析

**正員、工修

****学生員、工修

***正員、工博

*****学生員

東京大学大学院 工学系研究科 社会基盤工学専攻

(〒113-8656 文京区本郷 7-3-1

TEL : 03-3812-2111 ext.6128 FAX : 03-5689-7290)

効性を持たなくなるといった問題が生じることが知られている。また、OLS を用いた場合の分散推定値が偏りをもつため、通常の有意性検定は無効になってしまう。

(2) 空間的異質性

クロス・セクションデータに対して回帰分析などを行う場合、対象とする現象が空間上で構造的に安定していない場合（構造的不安定性 structural instability）がある。また、データの作成上の問題などのために誤差項の分散均一の仮定が崩れる場合（分散不均一性 heteroskedasticity）がある。Anselin ら¹⁾²⁾はそれらを総称して空間的異質性 Spatial Heterogeneity と呼んでいる。空間的異質性が存在する場合にも、通常の推定・検定手法を用いると、(1)と同様の問題が生じることが知られている。

(3) 可変地区単位問題

地域分析で取り扱われるデータは何らかの形で集計されているのが一般である。しかし、集計単位的设计が異なると分析手法は同じでも結果が大きく異なる場合がありうる事が古くから知られており、Openshaw ら⁴⁾はこれを可変地区単位問題 Modifiable Areal Unit Problem : MAUP と呼んでいる。MAUP はサイズ効果 scale effect とゾーニング効果 zoning effect の二つに大きく分類されるが、ある地域をいくつかの小地域に分割する場合、前者はその分割数が異なる場合に、後者は分割する地域数が同じであっても境界の設計が異なる場合に、それぞれ分析結果が一致しないことを指す³⁾。

ここにあげた(1)~(3)の問題はそれぞれ独立に生じる現象ではなく、相互に関連していることは言うまでもないが、本論文では(1)の空間的依存性、特にクロスセクションモデルの誤差項に空間的自己相関が引き起こされる場合のパラメータ推定法について着目する。

3. 空間的自己相関と統計学的対処法

(1) 空間的自己相関

以下、 y を被説明変数、 x_k ($k = 1, 2, \dots, m$) を説明

変数、 β_k ($k = 0, 1, 2, \dots, m$) をパラメータとする線形モデル(1)のパラメータ推定問題について考察する。

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m \quad \dots (1)$$

データの数を n ($> m+1$) とし、誤差項を含んだ統計モデルを以下のように行列表記する。

$$y = X\beta + \varepsilon \quad \dots (2)$$

$$\begin{aligned} y &= (y_1, \dots, y_i, \dots, y_n)^t, \\ \beta &= (\beta_0, \beta_1, \dots, \beta_m)^t, \quad X = \begin{pmatrix} 1 & x_{11} & \dots & x_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{mn} \end{pmatrix} \\ \varepsilon &= (\varepsilon_1, \dots, \varepsilon_i, \dots, \varepsilon_n)^t, \end{aligned}$$

ここで ε は誤差項であり、OLS の適用に際しては式(3)(4)を満たすものと仮定される。

$$E(\varepsilon) = 0 \quad \dots (3)$$

$$Var(\varepsilon) = \sigma^2 I \quad \dots (4)$$

σ^2 : 定数

空間を対象としたモデルの定式化にあたっては、明らかに空間依存性を引き起こすと考えられる要因はモデルに取り込まれていることが望ましいことは言うまでもない。しかし、実際にはそれらの要因をすべて考慮することは不可能であり、結果として残差に空間的自己相関が生じることが多い。

ところで、系列 series という言葉はもともと（通例3つ以上の）一続き・一組を意味するが、系列相関 serial correlation のうち統計学において最も盛んに研究が行われてきたのは言うまでもなく時系列 time series である。時間座標と空間的な座標を同一視すれば、時系列相関と空間的自己相関は本質的には同じであるとも考えられる。しかし、時系列相関に対しては、先行条件と現象との間の因果連鎖が時間の進行方向に副っており、ある時点における現象はその時点より過去へは影響を与えないという考えに基づきモデル化がなされる。これに対し空間的自己相関の場合には、3. (2) に示すように双方向に影響を与える構造としてモデル化されることが一般的であり、時系列とは異なるアプローチとも言えよう（無論、後者についても時間軸を導入した時空間モデルへと拡張し、時間軸上で一方向へ影響を与える構造とすることも可能である）。

たとえば、時系列モデルにおける系列相関の検出には Durbin-Watson 比が一般的に用いられるが、空

間的自己相関の検出にはそれと類似の Moran の I 統計量と呼ばれる指標が用いられる (I 検定)²⁾。Moran の I 統計量は、OLS を適用して得られた残差を用いて次式のように定義される。

$$I = \frac{N}{S} \frac{\varepsilon' W \varepsilon}{\varepsilon' \varepsilon} \quad \cdots (5)$$

ε : OLS による残差ベクトル

W : 空間重み行列 ($n \times n$)

N : データ数

$$S = \sum_j \sum_i w_{ij} \quad \cdots (6)$$

ここで、 W は空間重み行列 Spatial Weight Matrix と呼ばれるもので、地点、ゾーンあるいはメッシュ間の空間的依存の度合いを表す行列である。モデルを同定可能とするためには、分析者はこの W を先験的に与える必要があり、 W の各要素 w_{ij} をゾーン i とゾーン j の間の空間距離 d_{ij} を用いて次のように与えることが多い。

$$w_{ij} = c_j / d_{ij}^\alpha \quad (i \neq j), \quad w_{ii} = 0 \quad \cdots (7)$$

α : 定数 c_j : $\sum_i w_{ij} = 1$ とする規格化定数

各行要素の和を 1 に規格化すると $S = N$ となる。Moran の I 統計量を用いて定義される次式の Z は、観測数が十分に多いとき漸近的に正規分布 $N(0,1)$ に従うことが知れており、しかも計算が比較的容易であることから最も一般的に利用されている。

$$Z = \frac{I - E[I]}{\sqrt{\text{Var}[I]}} \quad \cdots (8)$$

$$E[I] = \text{tr}(MW) / (n - k) \quad \cdots (9)$$

$$\text{Var}[I] = \frac{\left\{ \text{tr}(MWMW') + \text{tr}(MW)^2 + [\text{tr}(MW)]^2 \right\}}{(n - k)(n - k + 2) - \{E[I]\}^2} \quad \cdots (10)$$

$$M = I - X(X'X)^{-1}X' \quad \cdots (11)$$

ここで、Moran の統計量 I の期待値及び分散に関する式(9)・(10)は、誤差項が正規分布するという仮定で導かれたものである²⁾。

(2) 空間的自己相関に対する統計学的対処法

これまで、空間計量経済学や計量地理学の分野で

空間的依存性が存在する場合のパラメータ推定および検定方法について様々な対処法が提案されてきた。以下に主なモデルを記す。

(i) mixed regressive-spatial autoregressive model²⁾

$$y = \rho W y + X\beta + \mu \quad \cdots (12)$$

(ii) mixed regressive-spatial crossregressive model¹⁰⁾

$$y = X\beta + WX\gamma + \mu \quad \cdots (13)$$

(iii) mixed regressive-spatial regressive model¹⁰⁾

$$y = \rho W y + X\beta + WX\gamma + \mu \quad \cdots (14)$$

(iv) regression model with autoregressive disturbances²⁾

$$y = X\beta + (I - \lambda W)^{-1} \mu \quad \cdots (15)$$

(v) mixed regressive-spatial autoregressive model with autoregressive disturbances²⁾

$$y = \rho W y + X\beta + (I - \lambda W)^{-1} \mu \quad \cdots (16)$$

(vi) spatial autoregressive moving average model⁷⁾

$$y = \rho W y + X\beta + (I + \lambda W) \mu \quad \cdots (17)$$

(vii) spatial correlation model¹¹⁾

$$y = \rho W y + X\beta + Wv + \mu \quad \cdots (18)$$

$\rho, \lambda, \gamma = (\gamma_1, \dots, \gamma_m)'$: パラメータ

$\mu = (\mu_1, \dots, \mu_n)'$, $v = (v_1, \dots, v_n)'$: 確率変数
 $\cdots (19)$

$$\text{Var}(\mu) = \sigma_\mu^2 I, \text{Var}(v) = \sigma_v^2 I, E(\mu v') = 0$$

σ_μ^2, σ_v^2 : 定数

パラメータの推定に最尤法を用いる場合には、次のような正規分布が仮定される。

$$\mu \sim N(0, \sigma_\mu^2 I), v \sim N(0, \sigma_v^2 I) \quad \cdots (20)$$

より拡張されたモデルも含め、既存のほとんどのモデルは以下に示す一般形で表現することが可能である。

$$y = \sum_k \rho_k W_k y + X\beta + \sum_h W_h X\gamma_h + \varepsilon \quad \cdots (21)$$

$$\varepsilon = \begin{cases} \sum_l \lambda_l W_l \varepsilon + \mu & \cdots (22) \\ \sum_l \lambda_l W_l \mu + \mu & \cdots (23) \\ \sum_l W_l v_l + \mu & \cdots (24) \end{cases}$$

$\rho_k, \lambda_l, \gamma_h = (\gamma_{h1}, \dots, \gamma_{hm})'$: パラメータ

$$\mu_i = (\mu_1, \dots, \mu_n)', v_i = (v_1, \dots, v_n)': \text{確率変数}$$

$$\mu_i \sim N(0, \sigma_\mu^2 I), v_i \sim N(0, \sigma_v^2 I), E(\mu v_i') = 0$$

これらのモデルを用いる際には、前述のように分析者の先験情報をもとに Spatial Weight Matrix W を特定化しなければならない。ところが、この W の特定化を誤ると推定・検定に関する議論のいくつかが成立しなくなる。Florax ら⁹⁾はモンテカルロ実験によって特定化を誤った場合の影響を調べており、データサンプル間の結びつきを真の結びつきより過大評価した場合、過小評価した場合に比べて分析の信頼性がより低下することを示している。このように、 W の特定には細心の注意が必要であるが¹²⁾、現時点ではその特定化に関して指針となるような方法は確立していない。そのため、たとえば式(7)において α をいろいろ変える等いくつかの重みを設定した上で、感度分析的な考察を行うといった作業も場合により必要となるであろう。

さきに述べた Moran 統計量を用いた I 検定は、OLS による残差の空間的自己相関を検定する方法ではあるが、(i)～(vii)のモデルにおける ρ や λ の有意性を直接検定するものではない¹³⁾。そのため、ラグランジュ乗数検定 (LM 検定) も I 検定とともに利用され、特に内生変数の自己回帰項の存在の診断に有効である²⁾。LM 検定は、尤度比検定やワルド検定とともに最尤法に基づいた検定であるが、OLS 残差をもとに検定可能である点において簡便である。また、 I 検定に比べると理論的にもより明確である。しかし、小標本特性については不明であり、また、誤差項の自己回帰モデルと移動平均モデルを区別できないという問題点もある²⁾。

Anselin ら^{5),7)}はモンテカルロ実験によって各種検定法の検定力を比較しており、 I 検定および LM 検定の有効性を確認している。 I 検定にせよ LM 検定にせよ、いずれも理論的にはまだ不十分な点を残しているため、最終的には各種検定結果をもとにした分析者の総合的な判断が要求される。

4. 実際のデータを用いた実証比較

(1) 実証比較のための手法の選定

誤差項に空間依存性が認められる場合には、3.で

あげたような様々な統計学的対処法が考えられる。これらの手法については、どのモデルが一番優れているとは一概には言えないため、分析者がそれらのうち一つか二つを選択的に適用する可能性も大きい。その際、適用する手法によって分析結果が大きく異なるのか否かということは、政策の影響分析を行う際の大きな関心事となる。

これまで、3.であげた(i)～(vii)のモデルのうち、2～3程度のモデルを比較分析した研究はいくつかあるが、多くは一般的特性を検討するための仮想シミュレーションデータを用いている^{5),6),7)}。本論文では、誤差項に空間的自己相関が認められる場合に対する対処法として、最も基本的な以下の5つの手法(a)～(e)について、実際の地域分析を想定した適用を行い、パラメータの推定結果やモデルの予測精度等にどの程度の違いが生じるかを検証する。

$$(a) \quad y = \rho W y + X \beta + \mu \quad \dots (25)$$

$$(b) \quad y = X \beta + W X \gamma + \mu \quad \dots (26)$$

$$(c) \quad y = X \beta + \varepsilon, \quad \varepsilon = \lambda W \varepsilon + \mu \quad \dots (27)$$

$$(d) \quad y = X \beta + \varepsilon, \quad \varepsilon = \lambda W \mu + \mu \quad \dots (28)$$

$$(e) \quad y = X \beta + \varepsilon, \quad \varepsilon = W v + \mu \quad \dots (29)$$

$$\mu \sim N(0, \sigma_\mu^2 I), v_i \sim N(0, \sigma_v^2 I), E(\mu v_i') = 0$$

また、これらとの比較に用いる OLS を前提とした基本モデルを以降(o)と記す。

$$(o) \quad y = X \beta + \mu \quad \dots (30)$$

なお、(c)～(d)のようなモデル化は、誤差修正モデル error correction model or error component model: ECM と呼ばれる。 $E(\mu v) = 0$ を仮定しているので、通常のパラメータ推定手法を用いるうえでは、これらは誤差の分散共分散行列を以下のように構造化したと考えても同じである。

$$V(\varepsilon) = \sigma_\mu^2 \left[(I - \lambda W)' (I - \lambda W) \right]^{-1} \quad \dots (31)$$

$$V(\varepsilon) = \sigma_\mu^2 \left[(I + \lambda W)' (I + \lambda W) \right] \quad \dots (32)$$

$$V(\varepsilon) = \sigma_v^2 W' W + \sigma_\mu^2 I \quad \dots (33)$$

(2) 実証に用いたモデルとデータセット

具体的なデータを用いた既存の実証研究では、地区の犯罪率⁸⁾や発病率⁹⁾といった例が多いものの、

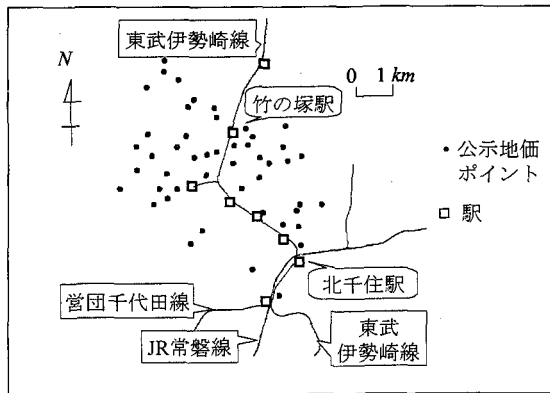


図-1 モデル適用対象地域の概略

社会基盤整備プロジェクトの影響分析において重要な要因となる地価のデータを用いた例としては、高塚・樋口^{13),15),16)}などその数は非常に限られている。そこで、本研究で公示地価を用いた回帰モデルを実証比較に用いることとする。具体的には、東京都足立区内の公示地価標準地のうち東武伊勢崎線北千住駅（ターミナル駅）から竹の塚駅までの各駅を最寄り駅とする合計 52 点を対象として、つぎのような回帰モデルを用いた。対象地域の概略は図-1 に示すとおりである。

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_4 x_{4i} + \varepsilon_i \quad \dots (34) \\ (i = 1, 2, \dots, 52)$$

y : 公示地価 (円/㎡)

x_1 : 標準地の地積 (㎡)

x_2 : 最寄り駅までの距離 (m)

x_3 : 最寄り駅から北千住駅までの所要時間 (分)

x_4 : 容積率 (%)

x_3 はWindows版「駅すばあと」((株)ヴァル研究所)を用いて算出した。それ以外のデータは、「地価マップ東京都 平成 8 年 ((財)土地情報センター編集)」による。

表-1 OLS によるパラメータの推定結果

	推定値	t値	標準偏差
β_0	3.12×10^5	18.5	1.69×10^4
β_1	2.03×10^2	2.79	72.8
β_2	-25.8	-6.27	4.12
β_3	-2.51×10^3	-2.09	1.20×10^3
β_4	2.31×10^2	5.16	44.7

通常最小二乗法によるパラメータの推定結果は、表-1 に示す通りである。自由度修正済みの相関係数は 0.897 であった。

(3) 空間的依存性の診断

図-2 は OLS による残差をプロットしたものである。図から、対象地域内北から北西部にかけて正の残差が、南部から南西部にかけて負の残差が、それぞれ偏って分布しており、空間的自己相関が生じている可能性が読み取れる。

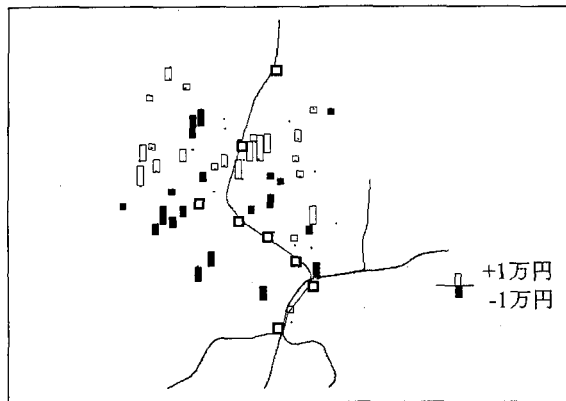


図-2 OLS による残差の分布

そこで、残差に対し 3.で述べた Moran の統計量を計算し、空間的自己相関の診断を行った。表-2 には、3つのパターンの Spatial Weight Matrix W についてこれらを計算した結果を示している。これらいずれのパターンについても 1%の有意水準で棄無仮説「空間的自己相関は存在しない」が棄却された。

表-2 Moran 統計量による空間的自己相関の診断

定数 α	0.5	1	2
Moran's I	0.062	0.195	0.013
Z	4.77	4.01	5.21
正規分布の上側確率	0	6.0×10^{-5}	0

正規分布の上側確率の欄に記載した 0 は、値が 4×10^{-6} 以下の極めて小さい値であることを示す。

(4) 各手法の適用結果

空間的自己相関が生じていると判断される場合、その原因となるような変数を見つけ、その追加・削

除あるいは変数変換等により空間的自己相関を解消することが可能であれば、必ずしも 3.に示したような対処法を用いる必要はあるまい。本研究でも、そのような可能性については検討を行ったが、パラメータの有意性等の観点において式(34)に替わるモデルを得ることはできなかった。そこで、空間的自己相関に対する対処法として、さきにあげた(a)～(e)の5つのモデルを適用した。なお、これ以降、 W に関しては式(7)における距離の減衰定数 α が2の場合

表－3 各モデルによるパラメータの推定結果

		推定モデル					
		(o)	(a)	(b)	(c)	(d)	(e)
AIC		1152	1149	1141	1146	1148	1149
ρ		—	0.48	—	—	—	—
λ		—	—	—	0.68	0.56	—
パラ メ ー タ	$\beta_0 (\times 10^5)$	3.12	1.45	2.24	3.17	3.21	3.03
	β_1	203	167	234	122	136	136
	β_2	-25.8	-18.5	-24.8	-28.2	-28.2	-26.1
	$\beta_3 (\times 10^3)$	-2.51	-1.76	-0.040	-1.49	-1.80	-1.04
	β_4	231	218	236	229	224	261
γ_1				909			
γ_2				32.2			
$\gamma_3 (\times 10^4)$				-15.1			

$\beta_0 (\times 10^5)$ は、たとえば、(o)についてはパラメータ β_0 の推定値が 3.11×10^5 であることを示す。本表の他の部分並びに以降の表についても同様。

表－4 標準誤差の推定値

		推定モデル					
標準誤差		(o)	(a)	(b)	(c)	(d)	(e)
$\sqrt{\sigma_\mu^2} (\times 10^4)$		2.10	1.93	1.83	1.86	2.01	0.16
$\sqrt{\sigma_v^2} (\times 10^4)$							2.5

表－5 推定パラメータの標準偏差

		推定モデル					
		(o)	(a)	(b)	(c)	(d)	(e)
パ ラ メ ー タ	$\beta_0 (\times 10^4)$	1.69	1.55	3.21	2.07	2.86	1.93
	β_1	72.8	66.7	64.6	63.0	66.5	59.4
	β_2	4.12	3.76	6.53	4.62	4.47	4.20
	$\beta_3 (\times 10^3)$	1.20	1.10	2.04	1.43	1.33	1.24
	β_4	44.7	41.0	40.2	44.5	45.1	42.6
γ_1				227			
γ_2				14.8			
$\gamma_3 (\times 10^3)$				4.42			

についてのみ結果を示すこととする。LM 検定等により各モデルのパラメータについての有意性についても検討したが、詳細はここでは省く。モデル(b)の WXy の部分については、有意であった x_1, x_2, x_3 のみを説明変数として用いている。

表-3～5 に、パラメータの推定値及びその標準偏差並びに標準誤差を示す。表-3 において、モデルによりパラメータの推定結果に割合大きな違いが認められる。(c)～(d)のような誤差修正モデルのパラメータ推定結果についても、OLS と比べて最大で5割以上の違いが認められ、また、この3手法相互においても最大3割程度の違いが認められる。一方、表-5 に示した標準偏差については、必ずしもすべてのパラメータについて一定の傾向が見られるわけではない。

OLS による結果と比べると、(a)～(e)のモデルはパラメータの数が増える分、基本的にはモデルの当てはまりは良くなる。そのことは、表-4 の標準誤差にも現れている。そこで、表-3 に示す各モデルについて赤池の情報量基準 AIC を計算した。AIC で判断する限りにおいては、(a)～(e)の各モデルはいずれも OLS を前提とした(o)に比べて有意に改善されていることがうかがえる。

(5) 各手法を用いた予測に関する考察

つぎに、モデルによって推定されたパラメータ・誤差分散を用いて、推定される地価とその標準誤差がどの程度異なるかを検討した。ここでは、ある1つの標準地についてのみ示すが、全体としてここに示した地点と同様に、推定値については実測値を基準に上下10%程度のばらつきが生じた。また、推定値の標準誤差については、(a)～(e)のモデルを用いることによって、多くの地点で(o)による標準誤差より小さくなる傾向が見られた。

表－6 地価の推定値とその標準誤差

標準地 足立 29	推定モデル						実測値
	(o)	(a)	(b)	(c)	(d)	(e)	
推定値 ($\times 10^5$)	3.78	3.88	3.99	3.74	3.77	3.79	4.31
標準誤差 ($\times 10^4$)	2.17	2.20	1.97	2.45	2.24	2.18	

最後に、各手法のパラメータ推定結果の差異が、実際の政策分析においてどの程度影響を与えるかを考察するために、仮想的なプロジェクトによる土地（住宅地）資産価値の上昇額を推定する場面を想定した試算を行った。竹の塚駅に急行が停車する等の交通改善により、対象地域の北側、面積にして 1/4 程の地域でターミナル駅である北千住駅までの時間が 3 分短縮されたとする。このとき、当該地域の住宅地資産価値の上昇額を、利子率 5% で地代換算した結果は、モデル(o)では 30 (±20) 億円/年程度であったものが、例えばモデル(c)では 20 (±15) 億円/年程度であった。これは、北千住駅までの時間に対応するパラメータ β_3 が、表-3 に示したようにモデル(o)と(c)では約 1.7 倍の違いがあり、その推定精度も異なるためである。同様に、他のモデルについてもパラメータの推定値・推定精度に応じた差異が生じることになる。ただし、対象地域内の住宅地面積は非常に大雑把な方法で求めたものであるため、ここでの概算値それ自体にはそれ程意味はないことを断っておく。しかし、説明変数の変化等に関して同一の条件下で計算を行っているため、概算値の相対的大小関係や比は一つの目安になると考えられる。なお、便益計測を目的とする場合には、small-open が成立しているのかといった理論的前提条件の詳細について吟味を要することは言うまでもない。

5. おわりに

本研究では、まず、回帰モデルの誤差項に空間的自己相関が存在する場合の統計学的対処法に関する既存研究のレビューと整理を行なった。

次に、それらのなかでも最も基本的な手法について、地価の推定を目的とした分析モデルへの適用を行い、その結果について考察した。無論、異なる手法を用いれば推定結果が異なるのは当然のことであり、結果の相異ということ自体が持つ意味は分析の目的に応じて違ってくる。

統計学における他の問題と同様、空間的自己相関への対処法についても、現時点においてはどの手法が一番優れているかという結論は得られていない。一方、分析者がこれらのなかからたまたま一つの手

法だけを用いた場合、ともすればそれによって得られた計算結果だけが実際の政策判断に用いられることも少なくない。しかし、本研究で用いた例では、政策に直接関わるパラメータの推定結果が採用する推定手法により大きく異なり、政策判断に重大な影響を与え得ることを示唆している。

もちろん、誤差項に空間的自己相関や分散不均一等が認められた場合には、まず、可能な限りデータの追加や加工、場合によっては関数形の変更等によりモデルを変更することも必要であり、本研究で扱ったような対処法は、あくまでそれらの手段を尽くしたのちに適用されるべきものであることを改めて強調しておきたい。

本研究の実証において得られた知見は、たかだか一組のデータセットによるということを考慮しても、社会資本整備の影響分析、とりわけ近年再びその重要性が認識されつつあるプロジェクトの便益評価における資産価値アプローチの適用等に対し、多くの示唆を与えるものとする。もちろん、空間計量モデルの統計学的取り扱い、Anselin¹⁸⁾も述べているとおり未だ unsatisfactory state of the art であり、分析の目的や分析者が先験的に有する情報に応じて適切に対処していかなければならないという側面を忘れてはなるまい。そのような意味においても、さらなる実証分析の蓄積が望まれる。

参考文献

- 1)Anselin, L. and Griffith, D.A.:Do spatial effects really matter in regression analysis?, *Papers of the Regional Science Association*, Vol.65, pp.11-34, 1988.
- 2)Anselin, L.:*Spatial Econometrics:Methods and Models*, Kluwer Academic, 1988.
- 3)Wong, D.W.S.: Aggregation effects in geo-referenced data, in *Practical Handbook of Spatial Statistics*, Arlinghaus, S.L., eds., CRC Press, pp.83-106, 1996.
- 4)Openshaw, S. and Taylor, P.J.: A million or so correlated coefficients: three experiment on the modifiable areal unit problem, *Statistical Applications in the Spatial Sciences*, Wrigley, N. and Bennett, R.J., eds., Pion, London, pp.127-144, 1979.
- 5)Anselin, L. and Rey, S.: Properties of tests for spatial dependence in linear regression models, *Geographical Analysis*, Vol.23, pp.112-131, 1991.
- 6)Florax, R.J.G.M. and Rey, S.: The impacts of misspecified

- spatial interaction in linear regression models, *New Directions in Spatial Econometrics*, Anselin, L. and Florax, R.J.G.M., eds., Springer, pp.111-135, 1995.
- 7)Anselin, L. and Florax, R.J.G.M.: Small sample properties of tests for spatial dependence in regression models, *New Directions in Spatial Econometrics*, Anselin, L. and Florax, R.J.G.M., eds., Springer, pp.21-74, 1995.
- 8)Getis, A. : Spatial filtering in a regression framework : examples using data on urban crime, regional inequality, and government expenditures, *New Directions in Spatial Econometrics*, Anselin, L. and Florax, R.J.G.M., eds., Springer, pp.172-185, 1995.
- 9)Haining, R. : *Spatial data analysis in the social and environmental sciences*, Cambridge University Press, 1990.
- 10)Florax, R.J.G.M. and Folmer, H. : Specification and estimation of spatial linear regression models, *Regional Science and Urban Economics*, Vol.22, pp.405-432, 1992.
- 11)Kelejian, H. and Robinson, D.: A Suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model, *Papers in Regional Science*, Vol.72, pp.297-312, 1993.
- 12)Griffith, D.A. : Some guidelines for specifying the geographic weights matrix contained in spatial statistical models, *Practical Handbook of Spatial Statistics*, Arlinghaus, S.L., eds., CRC Press, pp.65-82, 1996.
- 13)高塚創・樋口洋一郎: 空間的自己相関分析手法を用いた地価の空間的連関に関する統計的検証, *地域学研究*, Vol.26, No.1, pp.139-153, 1996.
- 14)蓼谷千鳳彦: 計量経済学の理論と応用, 日本評論社, 1996.
- 15)高塚創・樋口洋一郎: 期待の空間的依存性を考慮した地価モデル: その定式化と推定法, *応用地域学研究*, No. 2, pp.53-63, 1995.
- 16)樋口洋一郎・高塚創: 空間的自己相関の存在するデータが回帰分析に及ぼす影響に関する研究, *地域学研究*, Vol.25, No.1, pp.57-71, 1995.
- 17)Benirschaka, M. and Binkley, J. K. : Land price volatility in a geographically dispersed market, *American Journal of Agricultural Economics*, Vol. 76, pp.185-195, 1994.
- 18)Anselin, L.: Model validation in spatial econometrics: A review and evaluation of alternative approaches, *International Regional Science Review*, Vol.11, pp.279-316, 1988.

誤差項に空間的自己相関が存在する回帰モデルのパラメータ推定手法に関する考察

堤盛人・清水英範・福本潤也・井出裕史

空間計量モデル特有の問題である空間的依存性が存在する場合に、モデルのパラメータを通常の最小二乗法などの推定手法を用いて推定すると、分析結果の信頼性が低下するといった問題が生じる。本研究では線形回帰モデルを例に、誤差項に空間的自己相関が存在する場合のパラメータ推定手法について、これまで提案されている手法を整理した。次に、それらを公示地価を用いた簡単な地価モデルに適用して、パラメータ推定値と予測値ならびにそれらの精度について実証比較を行い、社会資本整備プロジェクトの影響分析上無視しえない結果の相異をもたらすことを示唆した。

A Study on the Statistical Inference Methods for Regression Models in the Presence of Spatial Autocorrelation

Morito TSUTSUMI・Eihan SHIMIZU・Jun-ya FUKUMOTO・Hiroshi IDE

Spatial autocorrelation of error terms in regression model is an important problem. In this paper, first, we review the suggested remedial methods for this problem and show their general form. Secondly, we apply these methods to a simple land price model and compare the estimates, prediction and the precision of the estimation empirically. As a result, it is demonstrated that parameter estimates are drastically changed in these methods and their differences can not be ignored in measuring the impact of infrastructure projects.
