

ベイズ推定を用いた巨視的土地利用予測モデル*1

Development of a Macroscopic Land Use Model using Bayesian Estimation Algorithm*1

和田沙織*2・菊池輝*3・岡田直也*4・北村隆一*5

By Saori Wada*2, Akira Kikuchi*3, Naoya Okada*4 and Ryuichi Kitamura*5

1. はじめに

交通施設整備あるいは土地利用規制や各種交通施策の実施は、周辺地域の土地利用状況に影響を与える。そのため、施設整備や各種施策が土地利用に及ぼす直接的・間接的影響を事前に把握した上で、その是非を判断する必要があり、これまで土地利用を予測するモデルが数多く構築されてきた¹⁾。

土地利用状況に与える影響は、上述の施設整備等によるものばかりではなく、当該地域の居住人口、サービス事業所数といったゾーン特性や経済動向といった要因に加え、ライフサイクルステージの変化や建物の老朽化などの個人特有の要因からも影響を受ける²⁾。加えて、ある時点の周辺地域特性が、時間を経、後の時点の土地利用状況に影響を与えることもある。このように多くの要因が複雑に関係し合っているため、これらをすべて考慮した土地利用モデルを構築することは、極めて困難であると言える。そこで、予測対象や目的に合わせて土地利用モデルが構築されてきた。

土地利用モデルは、予測対象の違いに着目すると、巨視的なモデルと微視的なモデルの2つに大きく分けられるが、本研究では、巨視的モデルのうち、土地利用間の相互作用や競合関係を表現する、連立方程式体系を用いた同時決定型モデルを対象とする³⁾。パラメータ推定にあたっては、多様なモデルに対応可能なベイズ推定によるパラメータ推定アルゴリズムを構築し、実データを用いた推定結果より、用途別土地利用間や、用途別土地利用と平均地価の関係を考察する。

2. 土地利用予測モデルの構造

(1) 同時方程式モデル

*1 キーワーズ：土地利用モデル、ベイズ推定

*2 正員、工修、日本ユニシス株式会社

*3 正員、工博、京都大学大学院工学研究科都市社会工学専攻

*4 学生員、京都大学大学院工学研究科都市社会工学専攻

*5 正員、Ph.D.、京都大学大学院工学研究科都市社会工学専攻

(京都市西京区京都大学桂4-C1-2、

TEL 075-383-3240, FAX 075-383-3236)

本研究では、土地利用間の相互作用を考慮するために、内生変数間の相互依存関係を記述可能な同時方程式モデルを採用する。 \mathbf{y}_i を内生変数ベクトル、 \mathbf{X}_i を外生変数行列、 \mathbf{B} と $\boldsymbol{\gamma}$ をそれぞれ係数行列、係数ベクトル、 $\boldsymbol{\varepsilon}_i$ を誤差項ベクトルとすると、同時方程式モデルは一般に、以下の式で表される。

$$\mathbf{y}_i = \mathbf{B}\mathbf{y}_i + \mathbf{X}_i\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_i \sim \text{i.i.d. } N(0, \boldsymbol{\Sigma}) \quad (1)$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1g} \\ \vdots & \ddots & \vdots \\ \sigma_{g1} & \cdots & \sigma_{gg} \end{pmatrix}$$

また、 g は方程式の数、 k は説明変数の数、 n をケースの数とし、 i は $1 \sim n$ の値をとるものとする。

また、各式の誤差項間には、同一ケース i において、以下のように相関があると仮定している。

$$E[\boldsymbol{\varepsilon}_{ai}, \boldsymbol{\varepsilon}_{bj}] = \begin{cases} \sigma_{ab} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

本研究では、京阪神都市圏パーソン・トリップ調査データ(PTデータ)に基づき、分析対象地域を京都府南部に設定し、予測単位をPTデータの入力ゾーン(計83ゾーン)とする。内生変数として、用途別土地利用面積変化量と平均地価を用い(表-1)、外生変数には、居住人口やサービス事業所数に加え、アクセシビリティや土地ポテンシャルを用いることとする。データの詳細については3.で述べる。

表-1 設定した内生変数

Y1	1980年~1990年住居系面積変化量
Y2	1980年~1990年商業系面積変化量
Y3	1980年~1990年工業系面積変化量
Y4	1990年平均地価
Y5	1990年~2000年住居系面積変化量
Y6	1990年~2000年商業系面積変化量
Y7	1990年~2000年工業系面積変化量
Y8	2000年平均地価

(2) ベイズ推定を用いたパラメータ推定方法

同時方程式モデルは、各式の誤差項間に相関があるため、方程式の右辺に含まれる内生変数と誤差項間に相関

が生じる。そのため、同時方程式モデルは、2段階最小二乗法 (2SLS) や3段階最小二乗法 (3SLS)、完全情報最尤法 (FIML) を用いてパラメータ推定が行われる。FIML同様、システム全体の尤度を最大化するという考えを用いながら、尤度そのものの計算が不要な推定手法としてベイズ推定がある。ベイズ推定は、多重積分などの計算を用いないため、計算の煩雑さが少ない、また、欠損値などの処理にも適しているという利点があり、計量経済や空間統計など多くの分野で用いられている。

さて、ベイズ推定手法を論じるにあたり、以降では簡単のため $\mathbf{B} = \mathbf{0}$ とし、右辺に内生変数を含まないSUR (Seemingly Unrelated Regressions) モデルであるとする。SURモデルは、FIMLを用いた場合、同時方程式モデルの推定と一致するという特徴がある。このとき、SURモデルの尤度関数は次式で表される。

$$f(\mathbf{y} | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \propto \prod_{i=1}^n |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\gamma})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\gamma})\right\} \quad (3)$$

ここで未知パラメータ $(\boldsymbol{\gamma}, \boldsymbol{\Sigma})$ の事前分布を、

$$\boldsymbol{\gamma} \sim N(\mathbf{c}_0, \mathbf{C}_0), \boldsymbol{\Sigma} \sim IW(n_0, \mathbf{S}_0) \quad (4)$$

とする。ここに、 \mathbf{c}_0 は $g \times 1$ のベクトル、 \mathbf{C}_0 は $g \times g$ の行列、 IW は (n_0, \mathbf{S}_0) をパラメータとする逆ウィッシュヤート分布を表し、 \mathbf{S}_0 は $g \times g$ の行列である。逆ウィッシュヤート分布の密度関数は、

$$h(\boldsymbol{\Sigma} | n_0, \mathbf{S}_0) \propto |\boldsymbol{\Sigma}|^{-\frac{n_0+g+1}{2}} \exp\left\{-\frac{1}{2} \text{tr}(\mathbf{S}_0^{-1} \boldsymbol{\Sigma}^{-1})\right\} \quad (5)$$

で与えられる。このとき、事後分布は、以下のように導かれる。

$$\boldsymbol{\gamma} | \boldsymbol{\Sigma}, \mathbf{y} \sim N(\mathbf{c}_1, \mathbf{C}_1), \boldsymbol{\Sigma} | \boldsymbol{\gamma}, \mathbf{y} \sim IW(n_1, \mathbf{S}_1) \quad (6)$$

$$\mathbf{c}_1 = \mathbf{C}_1 \left(\mathbf{C}_0^{-1} \mathbf{c}_0 + \sum_{i=1}^n \mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{y}_i \right)$$

$$\mathbf{C}_1^{-1} = \mathbf{C}_0^{-1} + \sum_{i=1}^n \mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{X}_i$$

$$n_1 = n_0 + n$$

$$\mathbf{S}_1^{-1} = \mathbf{S}_0^{-1} + \sum_{i=1}^n (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\gamma})(\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\gamma})'$$

以上のようにパラメータの事後分布を得ることができた。この事後分布から確率標本をサンプリングすることができれば、得られた標本の収束値を未知パラメータの推定値とみなせる。事後分布から確率標本を得る有効な方法として、マルコフ連鎖モンテカルロ法 (MCMC法) があるが、本研究ではMCMC法の1つであるギブ

ス・サンプリングを用いることとする。以下に式 (6) で与えられた事後分布に即して、ギブス・サンプリングを行う手順を示す。

- 1) 初期値を以下のように選定し、 $t=1$ とする。

$$(\boldsymbol{\gamma}^{(0)}, \boldsymbol{\Sigma}^{(0)}) = (\boldsymbol{\gamma}_{11}^{(0)}, \boldsymbol{\gamma}_{12}^{(0)}, \dots, \boldsymbol{\gamma}_{gk}^{(0)}; \boldsymbol{\sigma}_{11}^{(0)}, \boldsymbol{\sigma}_{12}^{(0)}, \dots, \boldsymbol{\sigma}_{gg}^{(0)})$$

- 2) $\boldsymbol{\Sigma}^{(t-1)}$ に基づいて、 $(\mathbf{c}_1^{(t)}, \mathbf{C}_1^{(t)})$ を次のように更新し、得られる多変量正規分布から $\boldsymbol{\gamma}^{(t)}$ をサンプリングする。

$$\mathbf{c}_1^{(t)} = \mathbf{C}_1^{(t)} \left(\mathbf{C}_0^{-1} \mathbf{c}_0 + \sum_{i=1}^n \mathbf{X}_i' \boldsymbol{\Sigma}^{(t-1)-1} \mathbf{y}_i \right) \quad (7)$$

$$\mathbf{C}_1^{(t)-1} = \mathbf{C}_0^{-1} + \sum_{i=1}^n \mathbf{X}_i' \boldsymbol{\Sigma}^{(t-1)-1} \mathbf{X}_i \quad (8)$$

- 3) 得られた $\boldsymbol{\gamma}^{(t)}$ を用いて、逆ウィッシュヤート分布のパラメータを次のように更新する。

$$\mathbf{S}_1^{(t)-1} = \mathbf{S}_0^{-1} + \sum_{i=1}^n (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\gamma}^{(t)})(\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\gamma}^{(t)})' \quad (9)$$

このとき $\boldsymbol{\Sigma}^{(t)}$ の密度関数は

$$h(\boldsymbol{\Sigma}^{(t)} | n_1, \mathbf{S}_1^{(t)}) \propto |\boldsymbol{\Sigma}|^{-\frac{n_1+g+1}{2}} \exp\left\{-\frac{1}{2} \text{tr}(\mathbf{S}_1^{(t)-1} \boldsymbol{\Sigma}^{(t)-1})\right\} \quad (10)$$

と与えられる。 $\boldsymbol{\Sigma}^{(t)}$ のサンプリングにあたっては、逆ウィッシュヤート分布ではなく、密度関数

$$h(\mathbf{Z}^{(t)} | n_1, \mathbf{S}_1^{(t)}) \propto |\boldsymbol{\Sigma}|^{-\frac{n_1-g-1}{2}} \exp\left\{-\frac{1}{2} \text{tr}(\mathbf{S}_1^{(t)} \mathbf{Z}^{(t)})\right\} \quad (11)$$

を持つウィッシュヤート分布からのサンプリングを行う⁴⁾。ここに、 $\mathbf{Z}^{(t)} = \boldsymbol{\Sigma}^{(t)-1}$ である。

- 4) $\boldsymbol{\gamma}^{(t)}$ 、 $\boldsymbol{\Sigma}^{(t)}$ の値を記録し、 $t=t+1$ とする。

2)から4)の手順を繰り返し、初期値の影響が十分に緩和された後、4)で記録した $(\boldsymbol{\gamma}^{(t)}, \boldsymbol{\Sigma}^{(t)})$ の平均値をもって、未知パラメータの推定値とする。

3. 使用データと分析指標

(1) 分析指標

土地利用変化は、即時的に表面化するものではなく、一定の時間を経て観測されるため、時間軸に沿ったデータが必要である。本研究では、1980年、1990年、2000年のPTデータを基準とし、これら年次における種々のデータを収集し、分析指標を作成した。具体的には、数値地図、国勢調査、事業所・企業統計調査、公示地価のデ

ータを用い、表-2に示す分析指標を作成した。

表-2 分析指標一覧

面積	住居系	最高地価 平均地価		
	商業系		変化量	最高地価
	工業系			平均地価
	その他都市的土地利用		人口	
	その他		世帯数	
開発可能面積	サービス事業所数			
ゾーンに占める 面積割合	住居系	変化量	人口	
	商業系		世帯数	
	工業系		サービス事業所数	
	その他都市的土地利用		従業員数	
	その他		密度	
開発可能面積	人口			
面積変化量	住居系	密度	世帯数	
	商業系		サービス事業所数	
	工業系		従業員数	
	その他都市的土地利用		アクセシビリティ	
	その他		土地ポテンシャル	
開発可能面積	ICまで最短平均所要時間			

(2) アクセシビリティ

交通アクセスの善し悪しは、土地利用形成要因の1つである。本研究では、交通のアクセスの善し悪しと空間相互作用を表現する指標として、次式で表されるアクセシビリティ指標を用いる。

$$A_i = \sum_{j=1}^n Y_j / t_{ij}^r \quad (12)$$

- A_i : ゾーン*i*のアクセシビリティ
- Y_j : ゾーン*j*の吸引指標
- t_{ij} : ゾーン*ij*間の空間的分離の指標
- n : ゾーン数
- r : 定数 (=2)

吸引指標には、ゾーンの居住人口、従業者数、工業系面積、集中交通量の4つを用い、またゾーン間の空間的分離の指標には、自動車平均所要時間と公共交通平均所要時間を用いた。これらを組み合わせ、合計8つのアクセシビリティ指標を作成した。

(3) 土地ポテンシャル

土地利用の形成には、立地主体者にとっての土地の魅力度が影響していると考えられる。そこで、立地時の土地の魅力度を表す指標を次式で定義する。

$$P_i = Z(carA_i) + Z(pubA_i) - Z(T_i) \quad (13)$$

$$carP_i = Z(carA_i) - Z(T_i) \quad (14)$$

$$pubP_i = Z(pubA_i) - Z(T_i) \quad (15)$$

- P_i : ゾーン*i*の土地ポテンシャル
- $carA_i$: ゾーン*i*の自動車アクセシビリティ
- $pubA_i$: ゾーン*i*の公共交通アクセシビリティ
- T_i : ゾーン*i*の平均地価
- $carP_i$: ゾーン*i*の自動車土地ポテンシャル
- $pubP_i$: ゾーン*i*の公共交通土地ポテンシャル
- Z : 変数の標準化

4. 推定結果

従来、土地利用の形成において、工業が他の土地利用を牽引していると考えられてきた。既存の土地利用モデルにおいても、工業立地が先決し、他の土地利用の立地は工業立地に従属して決定するという立地順序を仮定するモデルが見られる。しかし、1990年以降のわが国では工場の海外進出等による工業の衰退が見られる。その一方で、いわゆる都心部には商業の立地が多く見られることに加え、地価は商業地で最も高く、住宅地、工業地と続く傾向にあり、商業立地が主導となっているのではないかと考えられる。どちらの立地が主導であるかを判断することは困難であるため、工業主導の土地利用間の相互関係、商業主導の土地利用間の相互関係、それぞれの相互関係に基づく土地利用予測モデルを構築した。

本稿では紙面の制約上、内生変数間の相互関係が図-1のようになっている工業主導型モデルの一部の推定結果のみを掲載する(表-3)。

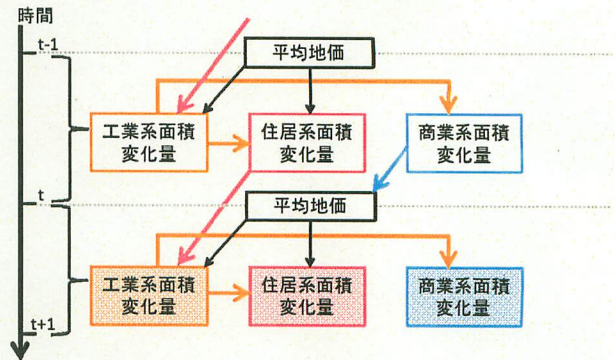


図-1 内生変数間の関係

(1) 工業系面積変化量

推定値を見ると、従業者の公共交通アクセス指標が負、平均地価の係数は正となっている。このことから、既に従業者にとって公共交通アクセスの良いゾーンほど工業系面積は増加しにくい傾向にあり、一方、地価が高いゾーンほど工業系面積が増加する傾向にあることがわかる。また、住居系面積変化量の係数が正となっていること、1980年~1990年の人口変化量の係数が正であること、1990年~2000年において、対象地域のほぼすべてのゾーンで工業系面積が減少していたことを総合すると、1980年~1990年に郊外化の進行により人口が減少していた都心部ほど、工業系面積が減少傾向にあることを示していると考えられる。

(2) 商業系面積変化量

土地ポテンシャルは正の係数を、公共交通土地ポテンシャルと人口の公共交通アクセス指標は負の係数を持っている。これは、公共交通のアクセスがよいゾーンでは

既に開発されているため新規立地が困難であるが、土地ポテンシャルの高い魅力的なゾーンには立地が進む傾向があると解釈できる。また、工業系面積変化量の係数は、経済状況に関わらず正となっており、工業立地と商業立地は共存する傾向にあることがわかる。

(3) 住居系面積変化量
 従業者の自動車アクセス指標とその他の面積変化量の係数は負であることから、従業者にとって自動車アクセスの良いゾーンでは住宅開発が既に行われているため、住宅系面積は新たに増加しにくい傾向にある一方、都市化の進行しているゾーンでは、住宅の新規立地が増加する傾向を示している。商業系面積の変化にも、工業系面積の変化は正の影響を及ぼしていることから、工業立地が増加すると、商業系、住居系の立地も増加し、工業が他の土地利用を牽引していることが伺える。

(4) 平均地価

従業者の公共交通アクセス指標、商業系面積割合の係数が正となっている。これは、わが国の地価は商業系土地利用で最も高いことから考えて、妥当な結果が得られている。一方、工業系の自動車アクセス指標は、工業への近接性の指標と考えられるが、係数は負となっており、工業面積が多いゾーン、もしくは周辺に工業立地の多いゾーンなどは、地価が低い傾向にあると解釈できる。

5. おわりに

土地利用間の相互依存関係を考慮するために、同時方程式モデルを用いた土地利用予測モデルを構築し、その推定には、多様なモデルに対応可能なベイズ推定を用いた。実データを用いた土地利用予測モデルのパラメータ推定を行った結果、工業立地が増加すると、商業系、

表-3 パラメータ γ の推定結果

1990~2000年住居系面積変化量	ベイズ推定					
	γ 平均値	標準誤差	95%信頼区間		Geweke統計量	確率
定数項	0.023	0.112	-0.200	0.241	-1.045	0.148
1990年従業者の自動車アクセス指標	-0.313	0.179	-0.667	0.036	4.682	1.000 *
1990年土地ポテンシャル	0.291	0.102	0.090	0.493	-1.434	0.076
1980年~1990年その他面積変化量	-0.381	0.119	-0.614	-0.146	-2.687	0.004 *
1990年~2000年工業系面積変化量	0.513	0.312	-0.199	1.047	4.748	1.000 *
1990年平均地価	0.231	0.176	-0.120	0.566	-1.660	0.048

1990~2000年商業系面積変化量	ベイズ推定					
	γ 平均値	標準誤差	95%信頼区間		Geweke統計量	確率
定数項	0.028	0.126	-0.225	0.278	-0.494	0.311
1990年土地ポテンシャル	0.676	0.144	0.392	0.956	-1.286	0.099
1990年公共交通土地ポテンシャル	-0.216	0.174	-0.563	0.126	1.490	0.932
1990年~2000年工業系面積変化量	0.468	0.421	-0.698	1.158	16.349	1.000 *
1990年人口の公共交通アクセス指標	-0.236	0.169	-0.552	0.103	7.773	1.000 *
1990年住居系面積割合	-0.279	0.102	-0.480	-0.081	4.348	1.000 *

1990~2000年工業系面積変化量	ベイズ推定					
	γ 平均値	標準誤差	95%信頼区間		Geweke統計量	確率
定数項	-0.002	0.121	-0.238	0.234	0.357	0.639
1990年従業者の公共交通アクセス指標	-0.211	0.181	-0.568	0.146	3.866	1.000 *
1980年~1990年人口変化量	0.132	0.110	-0.068	0.369	-2.252	0.012 *
1990年平均地価	0.168	0.186	-0.175	0.555	1.335	0.909
1990年工業系面積	-0.117	0.106	-0.346	0.075	8.188	1.000 *
1980年~1990年住居系面積変化量	0.366	0.275	-0.405	0.816	16.107	1.000 *

2000年平均地価	ベイズ推定					
	γ 平均値	標準誤差	95%信頼区間		Geweke統計量	確率
定数項	-0.087	0.022	-0.131	-0.045	-0.516	0.303
1990年平均地価	0.468	0.062	0.354	0.592	2.169	0.985 *
1990年従業者の公共交通アクセス指標	0.138	0.039	0.060	0.213	-3.397	0.000 *
1990年工業系の自動車アクセス指標	-0.042	0.024	-0.089	0.007	3.235	0.999 *
1990年商業系面積割合	0.026	0.045	-0.065	0.111	0.113	0.545
1990年~2000年商業系面積変化量	-0.023	0.030	-0.081	0.036	-0.009	0.496

住居系の立地も増加し、工業が他の土地利用を牽引していることが伺えるなど、土地利用間の相互依存関係を記述できたと言える。また紙面の制約上本稿には掲載できなかったが、用途別土地利用の誤差項間の相関を把握できたほか、単一の方程式ごとに重回帰分析を行った結果と比較して高い適合度が得られた。今後はパラメータの収束や事前分布パラメータの設定に関して、考察が必要である。

参考文献

- 1) 例えば；林良嗣，中村英夫，宮本和明，蔭山朝昭：都市近郊区土地利用変化予測モデル，土木計画学研究発表会講演集，Vol.1，1979.
- 2) 阿部宏史：広域都市圏における土地利用モデルの開発と応用に関する研究，京都大学学位論文，1985.
- 3) 同時決定型モデルの例としては；枝村俊郎，川井隆司：大都市圏の地目別土地利用モデルに関する研究，日本都市計画学会学術研究論文集，Vol.24，1989.
- 4) Liu, J.S.: *Monte Carlo Strategies in Scientific Computing*. Springer-Verlag, New York, 2001.