地価のヘドニックモデルを用いた便益評価への空間統計モデルの適用可能性*

A study on the application of spatial hedonic approach to project benefit evaluation *

堤盛人**・瀬谷創***

By Morito TSUTSUMI** • Hajime SEYA***

１．はじめに

不動産データは、言うまでもなく位置座標を持った空間データであり、データの位置に依存する空間的な依存性や異質性が存在する。近年、空間計量経済学・地球統計学と呼ばれる分野において、このような不動産データ相互の依存性や異質性を考慮したヘドニック価格関数の推定に関する研究が蓄積されてきた。しかしながら、これらの分野の手法が実際に社会資本整備の便益評価に用いられた例は筆者らの知る限り皆無であり、適用にあたっての問題点は必ずしも十分認識されてこなかった。

そこで本研究では、社会資本整備の便益評価において地価のヘドニック価格関数の推定にこれらの分野の手法を援用することとの理論的な問題を検討した上で、実際のデータを用いた実証分析を行う。

２．ヘドニック価格関数の推定

（１）基本モデル

本研究では、ヘドニック価格関数として次式のような線形回帰モデルを想定する。

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \,. \tag{1}$$

(1) 式において、$\boldsymbol{y}$ は$n\times 1$の従属変数ベクトル、$\boldsymbol{X}$ は$n\times k$ の説明変数行列、$\boldsymbol{\beta}$ は$k\times 1$のパラメータベクトル、$\boldsymbol{\varepsilon}$ は$n\times 1$の誤差項のベクトルである。本稿では (1) 式のモデルを基本モデル（basic model (BM)）と称す。誤差項に空間的自己相関が存在しながらBMのパラメータの推定にOLSを用いると、パラメータの有意性検定の際に t 値や F 値が過大に評価され、決定係数も過大になるといった問題が起こる。一方、空間計量経済学・地球統計学の分野では、それぞれ空間的自己相関をモデル化する手法が提案され、実証研究も蓄積されてきた。

（２）空間統計モデル

a）空間計量経済モデル

空間計量経済学の分野の空間的自己相関を考慮するモデルは、データ間の未知の依存関係を、従属変数同士の自己相関関係として捉えるspatial lag model (SLM) と誤差項同士の自己相関関係として捉えるspatial error model (SEM) に大別される。以下、本稿では、誤差項における空間的自己相関に着目するSEMを用いることとする。

SEMは、次式のように定式化される。

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}\,,\ \ \boldsymbol{\varepsilon} = \lambda \boldsymbol{W\varepsilon} + \boldsymbol{u} \,. \tag{2}$$

ここで、λは誤差項同士の影響関係を表す空間パラメータ、$\boldsymbol{W}$ は$n\times n$の空間重み行列、$\boldsymbol{u}$はその成分 u_i (i = 1, …, n) をiid誤差（パラメータを最尤法で推定する場合にはNID）とする$n\times 1$ベクトルである。空間重み行列$\boldsymbol{W}$の成分w_{ij} は、iとjに依存関係があるならば、0 でない何らかの値をとる。ここで、多くの場合$\boldsymbol{W}$は行和が1となるように行基準化される。SEMのパラメータ推定方法としては、最尤法、二段階最小二乗法、一般化モーメント法、ベイズ推定法等が提案されている。詳しくはLeSage and Pace (2008) 等を参照されたい。

b）地球統計モデル

地球統計学の分野の代表的なモデルである空間過程モデル（spatial process model (SPM)）は、データが連続的に分布する平面を前提としているという点にその特徴がある。具体的には、領域に次のような空間過程（ランダム場）を想定する。

$$\{\, y(\boldsymbol{s}) : \boldsymbol{s} \in D \subset \Re^2 \,\}, \tag{3}$$

$$\boldsymbol{s} = [c_x, c_y]' .$$

ここで、$\boldsymbol{s}$ は位置座標である。多くの場合、空間過程は次のような弱定常性を満たすと仮定される。

$$Cov[\,y(\boldsymbol{s}), y(\boldsymbol{s}+\boldsymbol{h})] = C(\boldsymbol{h}),\ \forall \boldsymbol{h} \in \Re^2 \,. \tag{4}$$

ここで、$C(\boldsymbol{h})$ は、共分散関数と呼ばれる$\boldsymbol{h}$のみに依存する関数である。すなわち、弱定常性の仮定とは、任意の2地点における確率変数間の共分散が、2地点間の距離と方位のみに依存するという仮定に他ならない。$C(\boldsymbol{h})$ が

*キーワーズ：便益評価、ヘドニック・アプローチ、空間統計学、空間計量経済学

**正会員、博(工)、筑波大学 大学院システム情報工学研究科（つくば市天王台1-1-1　tsutsumi@sk.tsukuba.ac.jp）

***非会員、修(社工)、株式会社パスコ

距離 $\|\boldsymbol{h}\|$ のみに依存する（すなわち方位には依存しない）とき、空間過程は等方的であるといわれる。SPMをヘドニック価格関数に用いる場合には、誤差項に弱定常性を仮定し、分散共分散行列を直接構造化することが多い。

$$E[\boldsymbol{\varepsilon}]=\boldsymbol{0}\,,\ Var[\boldsymbol{\varepsilon}]=\boldsymbol{\Sigma}(\boldsymbol{\theta})\,. \qquad (5)$$

ここで、分散共分散行列$\boldsymbol{\Sigma}$ は、その成分を共分散関数で与える$n\times n$ 行列である。共分散関数を特徴づけるパラメータベクトル$\boldsymbol{\theta}$ は、ナゲット、レンジ、シルという3つのパラメータからなり、これらの推定には、$\boldsymbol{\theta}$を重み付き最小二乗法 (WLS)、$\boldsymbol{\beta}$を実行可能な一般化最小二乗法 (FGLS) で繰り返し推定する方法や、制限付最尤法、ベイズ推定法等がある。詳細については、Cressie (1993)、Banerjee *et al.* (2004) 等を参照されたい。

ところで、不動産データを用いた分析を行う際には、単純化のため慣習的に等方的なコバリオグラムが用いられることが多いが、実際には等方性の仮定が満たされている場合はほとんどないであろう。共分散関数の異方性には、レンジのみが方向によって異なり、シルは方位に関係なく一定であるという幾何学的異方性と、両者とも異なるという帯状異方性の2種類がある（例えば、松田、小池 (2004)）。このうち幾何学的異方性については、次のような線形変換によって比較的簡単に考慮することが可能である。

$$\left[\boldsymbol{s}^*\right]=\begin{bmatrix}1 & 0\\ 0 & \delta\end{bmatrix}\begin{bmatrix}\cos(\delta) & \sin(\delta)\\ -\sin(\delta) & \cos(\delta)\end{bmatrix}\left[\boldsymbol{s}\right]. \qquad (6)$$

ここで、 φ は座標系の回転の角度を表すパラメータであり、δ は異方性比と呼ばれる、2方向のレンジの比を表すパラメータである。帯状異方性の考慮法は必ずしも確立されていないため、多くの場合、幾何学的異方性が存在するものとして処理されることが多い。

SPMは、領域全体に連続性と弱定常性を仮定しているため、自然な形で任意地点の値の内挿に用いることが可能である。この任意地点の値の内挿はクリギング（kriging）と呼ばれる。クリギングによる予測量は、予測誤差分散最小化により合理的に求められ、任意地点の値の最良線形不偏予測量（best linear unbiased predictor (BLUP)）を与える。クリギングによる地点$\boldsymbol{s}_0$における値の予測量は、次式により与えられる。

$$\hat{y}(\boldsymbol{s}_0)=x_0'\hat{\boldsymbol{\beta}}+\hat{\boldsymbol{c}}'\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})\,. \qquad (7)$$

ここで、$\boldsymbol{c}$ は観測値と予測値の共分散ベクトルである。予測誤差の分散（クリギング分散）は、 次式により求められる。

$$\sigma^2(\boldsymbol{s}_0)=C(\boldsymbol{0})-\hat{\boldsymbol{c}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{c}}+(\boldsymbol{x}_0-\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{c}})'(\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{x}_0-\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{c}})\,. \qquad (8)$$

３．便益評価と空間統計モデル

（１）一般的な便益算出過程

ある（small–openな）地域 A において地価データが図1の上の図のように分布しており、ある市場（例えば住宅市場）に関するヘドニック価格関数が (1) 式で与えられるとしよう。ここで、m 番目の不動産の特性を示す$n\times 1$ の列ベクトルを$\boldsymbol{x}_m$としたとき、$\boldsymbol{x}_m$の限界的な変化に対する不動産価格の限界的な変化（限界便益(marginal benefit)、or 限界評価額 (marginal implicit price)）は、

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}_m'}=\beta_m\,, \qquad (9)$$

で与えられる。これは、誤差項において空間効果を考慮するSEMやSPMであっても同様である。

次に、地域Aに効果計測を行うエリア区分（q = 1, …, Q）を設定し（図1下）、それぞれのエリア区分の代表点（中心点等）におけるプロジェクト前後における注目する説明変数（例えば、アクセシビリティ）の変化ベクトルΔx_m^q (q = 1, …, Q) を求める。Δx_m^q が求められれば、この値を限界便益に乗じ、さらにそれぞれのエリア区分における市場面積S^qを乗じれば、次式のように便益額が算出できる。

$$\sum_{q=1}^{Q}(\beta_m \Delta x_m^q S^q)\,. \qquad (10)$$

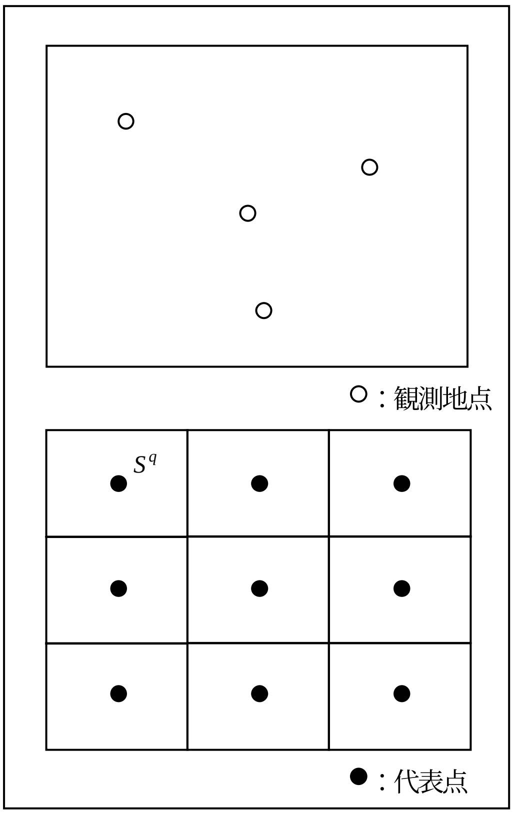

図１　代表点の設定

ところで、ヘドニック価格関数を推定する際には、データの特性から従属変数の不動産価格は次式のように対数変換されることが多い。

$$\ln \boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}\,. \tag{11}$$

(11) 式は次式と同値である。

$$\boldsymbol{y} = e^{\boldsymbol{X\beta}+\boldsymbol{\varepsilon}}\,. \tag{12}$$

この場合、簡単な計算から、$\boldsymbol{y}$ に対する限界便益は次式のように求められる。

$$\begin{aligned}\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}'_m} &= \beta_m e^{\boldsymbol{X\beta}+\boldsymbol{\varepsilon}} \\ &= \beta_m \boldsymbol{y}\,.\end{aligned} \tag{13}$$

限界便益の分析自体が興味の対象である場合には、$\boldsymbol{y}$ の標本平均を $\bar{y}$、標本標準偏差を $\bar{\sigma}_\varepsilon$ としたとき、$\bar{y}$ や、$\bar{y} \pm \bar{\sigma}_\varepsilon$ 等の比較点を設定し、 $\beta_m \bar{y}$、$\beta_m(\bar{y} \pm \bar{\sigma}_\varepsilon)$ のように限界便益を求めればよい。しかしながら、便益評価が興味の対象である場合は、$\boldsymbol{y}$ を標本モーメントで置き換えずに観測地点ごとに (13) 式で与えればよく、逆に上述のような置き換えは便益額の地域差を平滑化することになりかねない。

ここで、限界便益を (13) 式によって求める場合、設定する代表点における従属変数 y^q が未知であるため、代表点における限界便益が求められないという問題が生じる。この問題の解決策としては大きく次の2つ考えられる。まず、観測値平均によって限界便益を $\beta_m \bar{y}$ と与え、代表点においてもこの限界便益を用いるという方法である。この方法は簡便であるが、前述の通り便益額の地域差を平滑化してしまうという問題がある。もう1つの方法は、区分 q における説明変数ベクトルを $\boldsymbol{x}^q$ としたとき、$\ln \hat{y}^q = \boldsymbol{x}'^q \hat{\boldsymbol{\beta}}$ によって従属変数の対数をOLS内挿し、それを

$$\hat{y}^q = \exp(\boldsymbol{x}'^q \hat{\boldsymbol{\beta}})\,, \tag{14}$$

と変形するというものである。この方法は素朴で実用的であり、空間的自己相関が存在しない場合 $\ln \hat{y}^q$ は、BLUPとなるという利点がある。しかしながら、空間的自己相関が存在する場合、OLS予測量はBLUPとはならず、また、OLSによる予測値は観測値と一致しないため、(14) 式を用いた場合、観測地点において (13) 式が満たされないという問題が生じる。

したがっていずれの方法にも問題点があり、従属変数を対数変換した場合における、理論的に妥当でかつ実用的な限界便益算出方法の開発が必要であるといえる。以下では、これらの議論を踏まえた上で、地価関数の推定にSEM、SPMを用いることと、便益算出過程との整合性とについて議論する。

（2）便益算出過程とSEM

前述の通り、空間計量経済モデルは離散的な観測地点を前提としており、その地点における観測値間の空間相関関係を$\boldsymbol{W}$で記述するモデルである。したがって、観測点以外の地点の存在は前提としておらず、観測点とは別の地点を想定することは、$\boldsymbol{W}$の構造変化、すなわち空間的な均衡状態（spatial equilibrium）の崩壊を意味する。このような場合、$\boldsymbol{W}$の下で推定されたパラメータ推定値はもはや意味をなさなくなる。

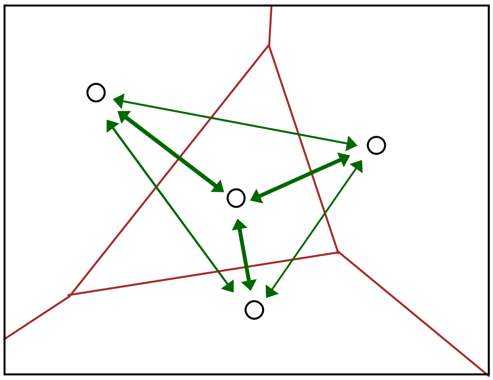

図２　空間計量経済モデルと便益評価

したがって、3–(1) で述べたような新たに代表点を設ける方法は、空間計量経済モデルの前提と整合しない。空間計量経済モデルの前提と整合的な便益評価を行うためには、観測地点間の空間的な均衡状態を維持したままで便益評価を行うことが必要である。このような目的を達成する最も素朴な方法として、例えば図2のように領域をボロノイ分割し、各領域内は均質であると仮定するという方法が考えられる。この方法であれば、新たに代表点を設ける必要が無いため、均衡状態における地点間の影響関係が保持される。また、従属変数を対数変換した際に、代表点における従属変数が必要になるという問題も生じない。無論、領域の分割方法はボロノイ分割に限られたものではなく、$\boldsymbol{W}$の構造変化を引き起こさない分割手法であれば、どのようなものでもよい。

（3）便益算出過程とSPM

地球統計学の分野のSPMは、データが連続的に分布する平面を前提としており、そもそも多くの場合、任意地点の値の内挿（クリギング）を目的として使用されるため、代表点の設定と不整合を起こさない。また、3–(1) で述べた、代表点を設定した場合の従属変数の対数変換に関する問題も解決できる。以下、これらについて

具体的に示そう。

従属変数に自然対数を取った場合、SPMを用いた場合の限界便益もまた (13) 式で与えられる。ここで、前述の通り新たに設けられた代表点 q における限界便益を求めるためには q における従属変数が必要となる。しかし、この値は自然なクリギング内挿によって、

$$\hat{z}^q = \ln \hat{y}^q = \boldsymbol{x}'^q \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{c}}' \hat{\boldsymbol{\Sigma}}^{-1} (\ln \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}),$$

$$\hat{y}^q = \exp(\hat{z}^q), \tag{15}$$

と求められる。(15) 式の観測地点における予測値は観測値と完全に一致するため、BMの場合（(14) 式）と違って観測地点で (13) 式が満たされるという利点がある。

ところで、(15) 式は、中央値予測量として知られるものである。(15) 式の予測量は、$\hat{y}^q$ に対して不偏とはならないため、不偏性を満たすように、バイアスを補正した予測量が提案されている。

$$\hat{y}^{*q} = \exp\left(\hat{z}^q + \left(\frac{1}{2}\right)\hat{\sigma}^2_{z,q} - \boldsymbol{m}'^q \boldsymbol{x}^q\right),$$

$$\boldsymbol{m}'^q = (\boldsymbol{x}^q - \boldsymbol{X}'\boldsymbol{C}^{-1}\boldsymbol{c})'(\boldsymbol{X}'\boldsymbol{C}^{-1}\boldsymbol{X})^{-1}, \tag{16}$$

ただし、$\hat{\sigma}^2_{z,q}$ は、代表点qにおけるクリギング分散である。(16) 式による予測は、lognormal kriging (LK) と呼ばれる。LKは、最も広く用いられているクリギング手法の一つであるが、分散を用いてバイアスを修正するため、予測値の感度が非常に鋭敏になるという問題がある。これに関して、Tolosana-Delgado and Pawlowsky-Glahn (2007) は、(15) 式のように、元データを対数変換して分析し、再び指数を取って元に戻す（back transform）方法の有効性を主張している。したがって本研究では、予測のために (15) 式を用いることとする。

４．実際のデータを用いた実証比較分析

以上の議論をもとに、瀬谷・堤(2007) と同じく、2005年に開通したつくばエクスプレス（TX）の沿線18市区町村を例に実証分析を行う。用いたデータは、1999年度の住宅用途の土地の公示地価データ1,074点である。

地価関数には、従属変数として地価（円/m^2）の自然対数を、説明変数としては表１に示す変数を用いる。

各モデルのパラメータの推定結果については紙面の都合上省略し、発表時に示すことする。便益額の試算結果は表２に示すとおりである。TXの総事業費が約8千3百億円であることを考えれば、これらの推計額は非常に大きく、過大推計の可能性がある。これは、そもそも対象地域が small とは呼べない程大きいこと等に起因するものと考えられる。しかしながら、本研究では便益額そのものは考察の対象外であるので、このことは無視する。

SAEMとSPMでは、両者ともBMに比べて３割近く小さい便益額を示した。SAEMと異方性を考慮した場合のSPMの便益額は非常に似通った値を示しており，本研究においては，空間的自己相関の考慮における誤差項の構造化手法（間接：SAEM，直接：SPM）は、便益額にそれほど大きな差異をもたらさなかった。

紙面の都合から、より詳細な考察については、発表時に示すこととする。

表1　説明変数

変数	変数名
切片（定数項）	切片
最寄駅までの距離（km）	駅距離
最寄駅から東京駅までの時間距離（時間）	東京駅時間
容積率（%）の自然対数	ln（容積）
つくば駅ダミー	つくば駅
水道ダミー	水道
ガスダミー	ガス
下水道ダミー	下水道
1低専ダミー	1低専
住居ダミー	住居
1中専ダミー	1中専
市街化調整区域ダミー	調整区域
常磐線ダミー	常磐線
常総線ダミー	常総線
龍ヶ崎線ダミー	龍ヶ崎線
北総線ダミー	北総線
成田線ダミー	成田線
野田線ダミー	野田線
伊勢崎線ダミー	伊勢崎線
埼玉高速鉄道ダミー	埼玉高速鉄道

表２　ボロノイ領域を用いた場合の
各モデルによる便益額（千億円）

	便益額
BM	41.04
SAEM	31.55
SPM（等方性仮定）	34.89
SPM（異方性考慮）	30.57

参考文献

瀬谷 創, 堤 盛人：地価から見たつくばエクスプレスの開発効果, 土木計画学研究・講演集, Vol.35, (CD-ROM講演番号: 118), 2007.

松田節郎，小池克明: 地熱地域における温度検層データを用いた地温分布の 3 次元解析，情報地質，15 (1), 15–24, 2004.

Banerjee, S., Carlin, B., and Gelfand, A. : *Hierarchical Modeling and Analysis for Spatial Data*, Chapman & Hall/CRC, 2004.

Cressie, N. : *Statistics for Spatial Data. Revised Edition*, John Wiley & Sons, 1993.

LeSage, J.P., and Pace, R.K. : *Introduction to Spatial Econometrics*, Crc Pr I Llc, 2008.

Tolosana-Delgado, R., and Pawlowsky-Glahn, V. : Kriging regionalized positive variables revised: Sample space and scale considerations, *Mathematical Geology*, 39, 529–558, 2007.