論　文

# 空間計量経済学とヘドニック価格アプローチ

加　藤　尚　史*
（名古屋大学大学院環境学研究科助教授）

## １．はじめに

計量経済モデルにおいて推定や予測を行うときには，撹乱項の自己相関を考慮する必要がある。注意すべきことは，横断面データを用いる場合であっても，撹乱項が関係し合うことはあり得るということであろう。空間自己相関と呼ばれるそうした関係を特定することは，推定や予測の精度を左右することになるので，慎重になされなければならない。空間計量経済学においては，特定方法に応じていろいろなモデルが提案されているけれども，ア・プリオリにモデルを選択することは難しいと考えられてきた。総体的にパフォーマンスの良いモデルを見つけることは，重要であるにもかかわらず，最近に至るまで試みられていなかったと見られる[1)]。

ヘドニック価格アプローチは，不動産の価格が特性に依存すると考えて，不動産市場を観察することで，特性に対する人々の評価を捉えようとするものである[2)]。価格と特性の関係は，ヘドニック価格関数と呼ばれる。空間計量経済学がこの関数を取り上げることは多い。いくつかの理由から空間自己相関が問題となり得るからである。

本論文の目的は，ヘドニック価格関数において，高い精度を有することが示されたモデルを使って撹乱項の自己相関を考慮したうえで，推定と予測を試みることにある。既存の論文は，アド・ホックにモデルを選んでいる点で，問題を含むと思われる[3)]。総体的にパフォーマンスの良いモデルを用いることは有用となろう。

次の節では，推定と予測の精度について最近なされた理論分析を紹介する。それらは，広く関心が持たれる線型のモデルにおいて一般化最小2乗法を使って推定と予測を行うという実験からなる。第3節では，理論分析の結果に基づいて，ヘドニック価格関数にかかわる応用分析を試みる。利用されるデータは，過去のものであるが，豊富な内容を有し，自己相関を考えるにあたって都合が良いものである。第4節では，今後の課題に言及する。

## ２．理論分析

線型モデルを $\mathbf{y} = \mathbf{Xb} + \mathbf{u}$ で表わすことにしよう。

---

* 1990年一橋大学大学院経済学研究科博士後期課程単位修得．金沢大学経済学部助教授，南山大学総合政策学部助教授などを経て，2002年より現職．専攻：計量経済学．所属学会：応用地域学会，日本経済学会，日本統計学会など．

1) Anselin and Bera（1998）や Anselin（2001）を参照．

2) 加藤（1996）を参照．

3) サーベイを目的とした論文の中で初めて空間自己相関に言及したのは，Pace et al.（1998）や Sheppard（1999）であると見られる．Palmquist（1999）は，自己相関を取り上げなかったが，Palmquist（2005）はそれを取り上げるようになった．これらは，空間自己相関が多くの論文において考慮されるようになってきていることを示唆する．

$$\mathbf{y}=\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \mathbf{y}_3\end{bmatrix},\ \mathbf{X}=\begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\\ \mathbf{X}_3\end{bmatrix},\ \mathbf{b}=\begin{bmatrix}b_0\\ \vdots\\ b_M\end{bmatrix},\ \mathbf{u}=\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\\ \mathbf{u}_3\end{bmatrix}$$

である。$[\mathbf{y}_1 \quad \mathbf{X}_1 \quad \mathbf{u}_1]$と$[\mathbf{y}_2 \quad \mathbf{X}_2 \quad \mathbf{u}_2]$は，それぞれ，推定と予測に使われる$N_1$組と$N_2$組のデータを示し，$[\mathbf{y}_3 \quad \mathbf{X}_3 \quad \mathbf{u}_3]$はそれ以外の$N_3$組のデータを示す。$\mathbf{b}$は$M$+1 個の係数パラメターを含んでいる。$\mathbf{u}$は，撹乱項を表わし，平均が$\mathbf{0}$で共分散が

$$\sigma^2\mathbf{K}(\mathbf{D};\mathbf{c})=\sigma^2\begin{bmatrix}\mathbf{K}(\mathbf{D};\mathbf{c})_{11} & \mathbf{K}(\mathbf{D};\mathbf{c})_{12} & \mathbf{K}(\mathbf{D};\mathbf{c})_{13}\\ \mathbf{K}(\mathbf{D};\mathbf{c})_{21} & \mathbf{K}(\mathbf{D};\mathbf{c})_{22} & \mathbf{K}(\mathbf{D};\mathbf{c})_{23}\\ \mathbf{K}(\mathbf{D};\mathbf{c})_{31} & \mathbf{K}(\mathbf{D};\mathbf{c})_{32} & \mathbf{K}(\mathbf{D};\mathbf{c})_{33}\end{bmatrix}$$

となる正規分布に従うと仮定する。$\sigma^2$は標準的なパラメターであり，$\mathbf{K}(\ \ ;\mathbf{c})$は同次数の行列をインプットとアウトプットにとる関数である。

$$\mathbf{D}=\begin{bmatrix}\mathbf{D}_{11} & \mathbf{D}_{12} & \mathbf{D}_{13}\\ \mathbf{D}_{21} & \mathbf{D}_{22} & \mathbf{D}_{23}\\ \mathbf{D}_{31} & \mathbf{D}_{32} & \mathbf{D}_{33}\end{bmatrix}$$

は，データが観察される地点の間の距離を成分とする対称行列を示す。$\mathbf{c}$ は空間自己相関を規定する 2 個のパラメターからなる。それらを$c_1$と$c_2$で表わすことにする。$\mathbf{K}(\mathbf{D};\mathbf{c})_{11}$や$\mathbf{D}_{11}$は$N_1\times N_1$，$\mathbf{K}(\mathbf{D};\mathbf{c})_{22}$や$\mathbf{D}_{22}$は$N_2\times N_2$，$\mathbf{K}(\mathbf{D};\mathbf{c})_{33}$や$\mathbf{D}_{33}$は$N_3\times N_3$の部分行列である。

$\mathbf{K}(\mathbf{D};\mathbf{c})=\mathbf{I}$にならない限り，通常最小 2 乗推定量を用いることは効率的でなく，一般化最小 2 乗推定量を用いることが期待される。対数尤度関数を集約して

$$L(\mathbf{c})=-\frac{N_1}{2}(\ln 2\pi-\ln N_1+1)-\frac{N_1}{2}\ln(\mathbf{y}_1'\mathbf{H}'\mathbf{K}(\mathbf{D};\mathbf{c})_{11}^{-1}\mathbf{H}\,\mathbf{y}_1)-\frac{1}{2}\ln\left|\mathbf{K}(\mathbf{D};\mathbf{c})_{11}\right| \tag{1}$$

とし，最尤推定量$\tilde{\mathbf{c}}=[\tilde{c}_1\ \tilde{c}_2]'$を求めて，$\mathbf{c}$の代わりに使うと，

$$\hat{\mathbf{b}}=(\mathbf{X}_1'\mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{11}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{11}^{-1}\mathbf{y}_1 \tag{2}$$

$$\hat{\sigma}^2=(\mathbf{y}_1-\mathbf{X}_1\hat{\mathbf{b}})'\mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{11}^{-1}(\mathbf{y}_1-\mathbf{X}_1\hat{\mathbf{b}})/(N_1-M-1) \tag{3}$$

によって$\mathbf{b}$と$\sigma^2$を推定することができる。$\mathbf{H}=\mathbf{I}-\mathbf{X}_1(\mathbf{X}_1'\mathbf{K}(\mathbf{D};\mathbf{c})_{11}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{K}(\mathbf{D};\mathbf{c})_{11}^{-1}$である．$\hat{\mathbf{b}}$の共分散は

$$\hat{\mathrm{V}}(\hat{\mathbf{b}})=\hat{\sigma}^2(\mathbf{X}_1'\mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{11}^{-1}\mathbf{X}_1)^{-1} \tag{4}$$

として推定される。一般化最小 2 乗推定量と組み合わせて用いることが期待されるのは一般化最小 2 乗予測量である。それは，

$$\hat{\mathbf{y}}_2 = \mathbf{X}_2\hat{\mathbf{b}} + \mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{21}\mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{11}^{-1}(\mathbf{y}_1 - \mathbf{X}_1\hat{\mathbf{b}}) \tag{5}$$

となり，$\mathbf{X}_2\mathbf{b}$ を推定する部分と $\mathbf{u}_2$ を推定する部分によって構成される。

$\mathbf{K}(\mathbf{D};\mathbf{c})$ を特定する方法は直接的なものと間接的なものに分けられる。地理統計学的（geostatistical）な方法は，前者に分類され，$\mathbf{K}(\mathbf{D};\mathbf{c})$ の第 $ij$ 成分を $K(D_{ij};\mathbf{c})$ とするものである。$K(\ \ ;\mathbf{c})$ は相関関数（correlation function）と呼ばれる。$D_{ij}$ は第 $i$ 地点と第 $j$ 地点の間の距離を表わしている。一方，ウェイト行列（weight matrix）を用いた方法は，後者に分類され，$\mathbf{K}(\mathbf{D};\mathbf{c})$ を $\left\{(\mathbf{I} - c_1\overline{\mathbf{W}}(\mathbf{D};c_2))'(\mathbf{I} - c_1\overline{\mathbf{W}}(\mathbf{D};c_2))\right\}^{-1}$ とするものである。$\overline{\mathbf{W}}(\mathbf{D};c_2)$ はウェイト行列をその行和で基準化したものである。ウェイト行列の第 $ij$ 成分 $W_{ij}$ は第 $i$ 地点と第 $j$ 地点の近接性を表わす。したがって，それは撹乱項が互いに及ぼす影響を加重して捉えようとする方法であるということがわかる。

相関関数やウェイト行列の定義を変えれば，いろいろなモデルを考えることが可能になる。推定や予測の結果がモデルによって異なるか否かを検討するために Monte Carlo 実験を行うことは有用となる。地点を定めて，$\sigma^2$ と $\mathbf{c}$ を与えると，$\mathbf{u}$ を抽出することができるので，$\mathbf{X}$ と $\mathbf{b}$ を与えると，$\mathbf{y}$ を計算することができる。そこで，各モデルを使ってデータを作成したうえですべてのモデルを使って推定や予測を行えば，モデルを比較することが可能となる。

Dubin（2003）は，相関関数として

$$K(D_{ij};\mathbf{c}) = \begin{cases} 1 & \text{if } D_{ij} = 0 \\ c_1 \exp(-D_{ij}/c_2) & \text{if } D_{ij} > 0 \end{cases}$$

$$K(D_{ij};\mathbf{c}) = \begin{cases} 1 & \text{if } D_{ij} = 0 \\ c_1 \exp(-(D_{ij}/c_2)^2) & \text{if } D_{ij} > 0 \end{cases}$$

$$K(D_{ij};\mathbf{c}) = \begin{cases} 1 & \text{if } D_{ij} = 0 \\ c_1(1 - 3D_{ij}/2c_2 + D_{ij}^3/2c_2^3) & \text{if } 0 < D_{ij} < c_2 \\ 0 & \text{if } D_{ij} \geq c_2 \end{cases}$$

を取り上げている一方，ウェイト行列として

$$W_{ij} = \begin{cases} 0 & \text{if } D_{ij} = 0 \\ 1 & \text{if } D_{ij} > 0 \text{ and } S(D_{ij}) \leq c_2 \\ 0 & \text{if } D_{ij} > 0 \text{ and } S(D_{ij}) > c_2 \end{cases}$$

$$W_{ij} = \begin{cases} 0 & \text{if } D_{ij} = 0 \\ 1/D_{ij}^{c_2} & \text{if } D_{ij} > 0 \end{cases}$$

$$W_{ij} = \begin{cases} 0 & \text{if } D_{ij} = 0 \\ 1 & \text{if } 0 < D_{ij} \leq c_2 \\ 0 & \text{if } D_{ij} > c_2 \end{cases}$$

$$W_{ij} = \begin{cases} 0 & \text{if } D_{ij} = 0 \\ \exp(-D_{ij}/c_2) & \text{if } D_{ij} > 0 \end{cases}$$

とするものを取り上げている。$S(D_{ij})$ は第 $i$ 地点とその他の地点の距離を昇ベキの順にソートしたときの $D_{ij}$ の順位を示す。3 つの相関関数を用いたモデルをそれぞれ GS1，GS2，GS3 と呼んで GS モデルと総称し，4 つのウェイト行列を用いたモデルをそれぞれ WM1，WM2，WM3，WM4 と呼んで WM モデルと総称することにしよう。

推定にあたっては，(1)式の $\mathbf{K}(\mathbf{D};\mathbf{c})_{11}$ と(2)式から(4)式の $\mathbf{K}(\mathbf{D};\tilde{\mathbf{c}})_{11}$ を $\mathbf{K}(\mathbf{D}_{11};\mathbf{c})$ と $\mathbf{K}(\mathbf{D}_{11};\tilde{\mathbf{c}})$ で置き換えたものを使いながら [4]，グリッド・サーチによって(1)式の最大化を試みている。また，予測にあたっては，GS モデルでは(5)式の $\mathbf{D}$ を

$$\dot{\mathbf{D}} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix}$$

で置き換えた

$$\hat{\mathbf{y}}_2 = \mathbf{X}_2\hat{\mathbf{b}} + \mathbf{K}(\dot{\mathbf{D}};\tilde{\mathbf{c}})_{21}\,\mathbf{K}(\dot{\mathbf{D}};\tilde{\mathbf{c}})_{11}^{-1}(\mathbf{y}_1 - \mathbf{X}_1\hat{\mathbf{b}}) \tag{6}$$

を用いているのに対して，WM モデルでは

$$\hat{\mathbf{y}}_2 = \mathbf{X}_2\hat{\mathbf{b}} + \tilde{c}_1\overline{\mathbf{W}}(\mathbf{D}_{21};\tilde{c}_2)(\mathbf{y}_1 - \mathbf{X}_1\hat{\mathbf{b}}) \tag{7}$$

を用いている。 $\mathbf{K}(\dot{\mathbf{D}};\tilde{\mathbf{c}})_{11}$ と $\mathbf{K}(\dot{\mathbf{D}};\tilde{\mathbf{c}})_{21}$ は $\mathbf{K}(\dot{\mathbf{D}};\tilde{\mathbf{c}})$ を構成する $N_1 \times N_1$ と $N_2 \times N_1$ の部分行列である。(7)式を使う理由としては，2 つの地点にかかわる撹乱項の共分散がすべての地点の間の距離に依存して $\mathbf{K}(\mathbf{D};\mathbf{c})_{21}$ を求めることはできないために，それが一般的に使われているということを挙げている。

Monte Carlo 実験においては，2 次元 Euclid 空間で，$c_1$ を 0.6 や 0.8 にして，1000 組のデータを作成し，推定と予測のために 200 組と 80 組のデータを抽出している。$c_2$ は，GS1 と GS2，GS3 で 20 とされている一方，WM1 と WM2，WM3，WM4 で 5 と 2.5，7，2 とされている。データの作成と抽出，推定と予測を 200 回繰り返して，パラメターの推定値やその標準誤差，実現値に対する予測値の平均平方誤差を要約することでモデルを比較し，推定においてはいずれのモデルも良いパフォーマンスを示すけれども，予測においては WM モデルに比べて GS モデルが良いパフォーマンスを示すと述べている。

Kato (2005)は，WM モデルでは，$\mathbf{K}(\mathbf{D};\mathbf{c})_{21}$ ばかりでなく $\mathbf{K}(\mathbf{D};\mathbf{c})_{11}$ も求めることができないとしたうえで，推定にあたって $\mathbf{K}(\mathbf{D}_{11};\mathbf{c})$ を利用する限り，予測にあたって，$\mathbf{K}(\dot{\mathbf{D}};\mathbf{c})$ や $\mathbf{K}(\ddot{\mathbf{D}}_k;\mathbf{c})$ を利用することは可能であると考え，(6)式や

$$\hat{\mathbf{y}}_2 = \mathbf{X}_2\hat{\mathbf{b}} + \begin{bmatrix} \mathbf{K}(\ddot{\mathbf{D}}_1;\tilde{\mathbf{c}})_{21}\,\mathbf{K}(\ddot{\mathbf{D}}_1;\tilde{\mathbf{c}})_{11}^{-1} \\ \vdots \\ \mathbf{K}(\ddot{\mathbf{D}}_{N_2};\tilde{\mathbf{c}})_{21}\,\mathbf{K}(\ddot{\mathbf{D}}_{N_2};\tilde{\mathbf{c}})_{11}^{-1} \end{bmatrix}(\mathbf{y}_1 - \mathbf{X}_1\hat{\mathbf{b}}) \tag{8}$$

を用いて相対的に精度が高いほうを選択するということを提案している。

[4] (3)式の $N_1 - M - 1$ を $N_1$ で置き換えているといった点に注意する必要はないと思われる。

表１　予測のパフォーマンス

| データ作成モデル | GS モデル 最上位 | GS モデル 最下位 | WM モデル 最上位 | WM モデル 最下位 |
|---|---|---|---|---|
| $c_1 = 0.6$ の場合 | | | | |
| GS1 | GS1 | GS2 | WM4 | WM2 |
| GS2 | GS2 | GS1 | WM4 | WM2 |
| GS3 | GS3 | GS1 | WM4 | WM3 |
| WM1 | GS3 | GS1 | WM2 | WM1 |
| WM2 | GS3 | GS2 | WM2 | WM1 |
| WM3 | GS2 | GS1 | WM4 | WM1 |
| WM4 | GS1 | GS2 | WM4 | WM1 |
| $c_1 = 0.8$ の場合 | | | | |
| GS1 | GS1 | GS2 | WM4 | WM3 |
| GS2 | GS2 | GS1 | WM4 | WM2 |
| GS3 | GS3 | GS1 | WM4 | WM3 |
| WM1 | GS1 | GS3 | WM4 | WM1 |
| WM2 | GS3 | GS2 | WM2 | WM1 |
| WM3 | GS2 | GS3 | WM4 | WM2 |
| WM4 | GS1 | GS3 | WM4 | WM3 |

$$\ddot{\mathbf{D}}_k = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12,k} \\ \mathbf{D}_{21,k} & 0 \end{bmatrix}$$

であり，$\mathbf{D}_{12,k}$ と $\mathbf{D}_{21,k}$ は $\mathbf{D}_{12}$ の第 $k$ 列と $\mathbf{D}_{21}$ の第 $k$ 行である。$\mathbf{K}(\ddot{\mathbf{D}}_k;\tilde{\mathbf{c}})_{11}$ と $\mathbf{K}(\ddot{\mathbf{D}}_k;\tilde{\mathbf{c}})_{21}$ はそれぞれ，$\mathbf{K}(\ddot{\mathbf{D}}_k;\tilde{\mathbf{c}})$ を構成する $N_1 \times N_1$ と $1 \times N_1$ の部分行列を表わす。(6)式は，$N_1 + N_2$ 地点にかかわる撹乱項の共分散を $\sigma^2\mathbf{K}(\dot{\mathbf{D}};\tilde{\mathbf{c}})$ によって推定して，$\mathbf{y}_2$ の成分を同時に予測しようとするものであるのに対し，(8)式は，$N_1 + 1$ 地点にかかわる撹乱項の共分散を $\sigma^2\mathbf{K}(\ddot{\mathbf{D}}_k;\tilde{\mathbf{c}})$ によって推定して，それらを個別に予測しようとするものである。これらと(7)式は $\mathbf{u}_2$ を推定する部分において異なる。Dubin (2003) と同様にして実験を行うことで，WM モデルに基づいてデータが作成され，かつ，そのモデルが使われる場合には，より多くの地点を考慮したほうが良いので，(6)式が(8)式に優るということを裏付けている一方，それ以外の場合には，どちらが優るかはモデルによって異なるということを説明している。ただし，相対的な優位性は僅かなものにとどまるとも述べている。

表１は，予測におけるモデルのパフォーマンスを示したものである。Dubin (2003) に基づけば，WM2 によってデータが作成される場合にのみ，GS モデルの中で最下位にランクされたモデルは WM モデルの中で最上位にランクされたモデルに劣ることとなるのに対して，Kato (2005) に基づけば，少なくとも WM モデルによってデータが作成される場合には，WM モデルの中で最上位にランクされたモデルが，GS モデルの中で最下位にランクされたモデルに優り，最上位にランクされたモデルに優ることもあり得るということになる。これは，WM モデルにおいては(7)式に比べて(6)式や(8)式を用いたほうが良く，予測においては WM モデルに比べて GS モデルが良いパフォーマンスを示すとは言えないことを意味する。

Kato (2006) は，$\sigma^2$ と $\mathbf{b}$ を変えることで，追加的な実験を行っている。パラメターの変更は，特に設定を一般的なものとするために行われていて，$\mathbf{c}$ と $\mathbf{b}$，$\sigma^2$ の推定や $\mathbf{y}_2$ の予測について新たな知見を与える

ことにはならない。$\mathbf{y}=[\ln z_1 \ldots \ln z_{N_1+N_2+N_3}]'$ として，関心を $z_{N_1+k}$ の予測に向け，

$$\hat{z}_{N_1+k}=\exp(\hat{y}_{N_1+k}+\tfrac{1}{2}\hat{\sigma}^2_{N_1+k}) \tag{9}$$

という予測量を提案したうえで，

$$\hat{z}_{N_1+k}=\exp(\hat{y}_{N_1+k}) \tag{10}$$

と比較して，前者が後者に優ることを示している。$\hat{y}_{N_1+k}$ は $y_2$ の第 $k$ 成分の予測量であり，$\hat{\sigma}^2_{N_1+k}$ は $\sigma^2\mathbf{K}(\mathbf{D};\mathbf{c})$ の第 $N_1+k$ 対角成分の推定量である。(10)式は，撹乱項の形状にかかわりなく，広く使われているが，バイアスを伴う。(9)式は，$\mathbf{K}(\mathbf{D};\mathbf{c})=\mathbf{I}$ と見なすときに使うことが期待される

$$\hat{z}_{N_1+k}=\exp(\hat{y}_{N_1+k}+\tfrac{1}{2}\hat{\sigma}^2) \tag{11}$$

に呼応し，バイアスを削減しようとするものである。GS モデルにおいては，(9)式や(10)式は，いずれも，(6)式と組み合わされる。WM モデルにおいては，Dubin（2003）の見方に従えば，(7)式と(10)式を組み合わせざるを得ないのに対して，Kato (2005) の見方に従えば，(6)式や(8)式と(9)式を組み合わせることができる。したがって，実験の結果が予想されたものになっているということは理解されるであろう。

## ３．応用分析

名古屋市総務局企画部企画課・市民局地域振興部地域振興課の『学区別生活環境調査報告書』（昭和 61 年 12 月）は，小学校区を市民の日常生活圏として捉えたうえで，地図や指標を用いて，環境にかかわる学区ごとの特性と問題点を明らかにしようとしたものである。小学校区の数は 250 に及ぶ。調査時点は原則として昭和 60 年 4 月 1 日現在とされている。各学区の「生活環境図」には，昭和 60 年 1 月 1 日時点で国が定めた標準地とその公示価格が記されている。国土庁土地鑑定委員会の『地価公示』（昭和 60 年）を使って標準地の価格や構造特性に関する情報を獲得し，環境特性についてそれを含む小学校区の指標を『学区別生活環境調査報告書』から拾うことにしよう。ただし，『地価公示』に記載された標準地のうちのいずれが「生活環境図」に記されているかはわからないことから，名古屋市計画局土地対策課の『名古屋市地価公示・基準地位置図』（昭和 60 年 10 月）を使ってマッチングを行う。

名古屋市内には総じて 368 の標準地が分布しているが，「利用の現況」が「住宅」となっているものに焦点を絞ることにしたい。該当するものは，中区を除く 15 の区に属し，第 1 種住居専用地域や第 2 種住居専用地域，住居地域に位置する 229 の住宅地と準工業地域に含まれる 5 つの準工業地に分けることができる。後者は，商業地や工業地と同様に産業用地に属して前者と異なり，別な市場において取引される可能性が高いと思われるので，省きたい。環境指標の中には，調査時以後学区に変動があったにもかかわらず調査当時の小学校区の値が与えられているものや，学区によって値の得られないものが見受けられる。それらのうちのいくつかは分析を加えるにあたって無視し難い指標である。そこで，変動があったり欠測値が見られたりする小学校区とそこに位置する標準地は考慮せず，141 の学区に含まれる 214 の標準地を用いることとする。

表 2　被説明変数と説明変数

| 変　数 | 内　容 |
|---|---|
| 被説明変数 | |
| PRICE(円) | 価格に地積と割引率を掛けたもの. |
| 説明変数 | |
| 構造特性 | |
| SIZE($m^2$) | 一部が私道となっている場合にはその部分を含めた地積. |
| SHAPES | 台形や不整形であれば 1 とし四角形であれば 0 とするダミー変数. |
| DIRECTIONS | 前面道路の方位が南であれば 1 としそれ以外は 0 とするダミー変数. |
| WIDTH(m) | 前面道路の幅員. |
| PRIVATE | 前面道路が私道であれば 1 としそれ以外は 0 とするダミー変数. |
| GAS | ガス事業法による一般ガス事業または簡易ガス事業によりガスが供給されている場合，および，通常の工事費負担によってそれらからのガス供給が可能な場合には 1 としそれ以外では 0 とするダミー変数. |
| SEWERAGE | 下水道法の処理区域内にある場合，および，公共下水道に接続したり終末処理場があったりする場合には 1 としそれ以外では 0 とするダミー変数. |
| VOLUME(%) | 容積率. |
| 環境特性 | |
| PLAYGROUNDS($m^2$/人) | 児童公園や近隣・地区公園，どんぐり広場，児童遊園地の面積を人口で割ったもの. |
| SCHOOLYARDS($m^2$/人) | 児童 1 人当たりの運動場面積. |
| DOCTORS(人/千人) | 千人当たりの病院・診療所医師数. |
| DENTISTS(人/千人) | 千人当たりの病院・診療所歯科医師数. |
| FIRES(%) | 学区面積に対する過去 3 年間の建物焼損面積合計の割合. |
| ACCIDENTS(件/ha) | 昭和 59 年の 1ha 当たり人身事故件数. |
| POLLUTION(件/ha) | 昭和59 年度の公害指導件数を学区面積で割ったもの. |
| PARKSETC(%) | 公園・緑地と宗教・文化用地の割合. |
| WOODS(%) | 100$m^2$以上の一団をなす樹林地の昭和 58 年度における割合. |
| MEETING(箇所/千人) | 千人当たりの集会施設数. |
| BARSETC(箇所/100ha) | 商業統計調査による飲食店 2 類(バー，キャバレー，ナイトクラブ，酒場，小料理屋)に該当する風俗営業施設の 100ha 当たりの数. |
| RESIDENTIAL(%) | 住居用地の割合. |
| INDUSTRIAL(%) | 工業用地の割合. |
| UNDERUSED(%) | 未利用地(空閑地，駐車場，材料置場，農地，山林等)の割合. |
| PAVED(%) | 道路舗装率. |
| TIME(分) | 小学校から都心の栄までの所要時分. |
| STORES(箇所/千人) | 千人当たりの飲食店を除く小売業商店の数. |
| RESTAURANTS(箇所/千人) | 千人当たりの風俗営業施設を除く飲食店の数. |

取り上げた変数は表 2 に示されている。一般に，土地の取引は分割可能性に乏しくロット単位で行われ，ヘドニック価格関数は特定な期間を対象にして組まれたモデルにおいて導き出されるために，1$m^2$当たりの公示価格に地積と割引率を掛けることによって被説明変数を定義したい。ここでは，0.004 を使って月当たりの値とする [5]。建物の構造に関する情報を得ることは可能であるが，更地としての価格が公示されることになっているので，それらを考慮することは必要とされない。公害指導件数については，原データの形で表示されていたけれども，特に空間的な分布を扱うため，面積で割ったものを用いる。

データに記載ミスがあると疑われるときには，関係機関に問い合わせたうえで，ミスを訂正したりした。これは，標準地が近隣地域の中で標準的な画地を指していることや取り上げる標準地を絞っていることと合わせて，外れ値にかかわる問題を生じさせない理由となる。

[5] 年率に換算すると，約 5%になる．森杉（1991）によれば，割引率に 4%ないし 6%を当てることが多いとされている．

ヘドニック価格関数においては，被説明変数は一般に対数変換されるので，Monte Carlo 実験の結果に配慮しながら推定と予測を行うことができる。よく用いられる log-linear 型を取り上げることにしよう。ただし，説明変数が 0 を観測値に有する際には，対数変換を施すことは許されず，困難が生じる。定数を足して底上げするということは，恣意性を伴うことから，避けたほうが良いと判断した。そこで，PLAYGROUNDS，DOCTORS，DENTISTS，FIRES や POLLUTION，MEETING，BARSETC，INDUSTRIAL に関しては，ダミー変数と同じように，変換しないまま組み入れることとする。

標準地の価格が判定される場合には，取引事例との比較が行われることが多い。一方，環境を測定するために小学校区を利用することは，誤差を含むかもしれない。これらは，考慮すべき特性が省かれていないと見なしても，空間自己相関を招く可能性があると考えられる。既存の論文では，主に GS1 や WM1，WM2，WM3 が用いられてきたが，この論文では，GS3 と WM4 を取り上げることにしたい。実験においてそれらは GS モデルと WM モデルの中で総体的に最も良いパフォーマンスを示していたからである。標準地間の距離に関しては，『名古屋市地価公示・基準地位置図』を使って㎞単位で測定する。

空間自己相関を扱う場合，標準的なソフトウェアは，専用のコマンドを備えていないものの，行列計算と非線型最適化を可能にするものであれば，有用となる。この論文では，*TSP*（*Time Series Processor*）を用いることにしよう[6]。最尤推定を行うにあたっては，ML コマンドを使うことが考えられる。Version 4.4 以上においては，これに合わせて 2 通りの方法が提供されている。FRML は，尤度関数を集約しようとする際には適当でない。PROC は，パラメターの推定値の分散を最も小さくすると思われる Newton 法を取り上げることができないという点で制約を有する。quasi-Newton 法としては，BFGS（Broyden-Fletcher-Goldfarb-Shanno）法や DFP（Davidon-Fletcher-Powell）法が用意されている。前者は，後者に優ると考えられているが[7]，最適化を保証するものではない。BFGS 法に限らず必要に応じて DFP 法を取り上げることにしよう。パラメターの初期値を獲得するために，グリッド・サーチを試みる。Monte Carlo 実験は一般的な設定のもとで行われるということを考慮して，理論分析に基づき，GS3 においては $c_1 \in \{0.1, 0.2, \cdots, 1.0\}$，$c_2 \in \{10, 15, \cdots, 50\}$ と見なし，WM4 においては $c_1 \in \{0.1, 0.2, \cdots, 0.9\}$，$c_2 \in \{1.0, 1.5, \cdots, 4.0\}$ と見なしてサーチを行った後に，ステップの幅を半減させることによってサーチを 3 回繰り返すことにする。もっとも，尤度が増加する方向に対象を拡げてピークに至ることを確認しなければならなくなるかもしれない[8]。注意すべきことは，理論分析においては，$\mathbf{c}$ が既知であり，関心は $\tilde{\mathbf{c}}$ に向けられていて，グリッド・サーチの目的がパラメターの最終値を獲得することにあったということであろう。応用分析においては，$\mathbf{c}$ は未知であり，それに関して検定を加えることが必要とされる。ここでは，アルゴリズムを利用してパラメターの最終値を獲得するとともに，2 次導関数を数値計算して $\tilde{\mathbf{c}}$ の共分散を推定する。

表 3 は，一般化最小 2 乗推定の結果を与えている。モデル 1 は $\mathbf{K}(\mathbf{D}; \mathbf{c}) = \mathbf{I}$ と仮定したものである。モデル 2 と 3 は GS3 と WM4 に相当する。いずれにおいても $\mathbf{c}$ は DFP 法によって推定されている。一般に，モデルを比べる場合には，同一のアルゴリズムを用いることが求められる。BFGS 法は最適化を達成しないことがあった。$\tilde{c}_1$ の値は，モデルを問わず，有意である。これは，強い正の自己相関があり，近接する標準地について撹乱項の値が同じような方向に分布することを示す。また，それは，モデル 2 において，1 と有意に異なっていて，標準地が僅かにでも離れると，撹乱項の相関は弱まるというナゲット効果

[6] 空間的な統計解析のためのモジュールを持つものについては，Dubin et al.（1999）を参照.
[7] アルゴリズムを特に指定しない場合，BFGS 法が選択されることになる.
[8] $C_1$ は，GS モデルでは 1 以下となり，WM モデルでは 1 未満となることが求められる.

(nugget effect)が存在していることを意味する。

表3　ヘドニック価格関数の推定

| | モデル1 | モデル2 | モデル3 |
|---|---|---|---|
| $b_0$ | 4.3824*** (0.8365) | 5.5689*** (0.7543) | 5.3377*** (0.7463) |
| $b_{\mathrm{SIZE}}$ | 1.0995*** (0.0185) | 1.0604*** (0.0130) | 1.0581*** (0.0137) |
| $b_{\mathrm{SHAPES}}$ | −0.0272 (0.0311) | −0.0078 (0.0212) | −0.0160 (0.0212) |
| $b_{\mathrm{DIRECTIONS}}$ | 0.0320** (0.0166) | 0.0145 (0.0118) | 0.0134 (0.0118) |
| $b_{\mathrm{WIDTH}}$ | 0.0727** (0.0347) | 0.0660*** (0.0250) | 0.0642*** (0.0258) |
| $b_{\mathrm{PRIVATE}}$ | −0.2112*** (0.0859) | −0.1418*** (0.0601) | −0.1326** (0.0629) |
| $b_{\mathrm{GAS}}$ | 0.2510*** (0.0578) | 0.1575*** (0.0492) | 0.1237*** (0.0519) |
| $b_{\mathrm{SEWERAGE}}$ | 0.0823*** (0.0226) | 0.1159*** (0.0214) | 0.1131*** (0.0219) |
| $b_{\mathrm{VOLUME}}$ | −0.0174 (0.0237) | 0.0042 (0.0179) | 0.0069 (0.0178) |
| $b_{\mathrm{PLAYGROUNDS}}$ | 0.0085*** (0.0034) | 0.0078*** (0.0030) | 0.0082*** (0.0033) |
| $b_{\mathrm{SCHOOLYARDS}}$ | −0.0544*** (0.0192) | −0.0275** (0.0148) | −0.0317** (0.0154) |
| $b_{\mathrm{DOCTORS}}$ | 0.0040* (0.0026) | −0.0002 (0.0021) | 0.0009 (0.0022) |
| $b_{\mathrm{DENTISTS}}$ | −0.0020 (0.0091) | 0.0031 (0.0075) | 0.0022 (0.0080) |
| $b_{\mathrm{FIRES}}$ | 0.2387 (0.4750) | −0.1073 (0.3837) | 0.0089 (0.3796) |
| $b_{\mathrm{ACCIDENTS}}$ | 0.0154 (0.0255) | 0.0205 (0.0200) | 0.0261 (0.0205) |
| $b_{\mathrm{POLLUTION}}$ | −0.0101 (0.2226) | 0.1145 (0.1705) | 0.1008 (0.1769) |
| $b_{\mathrm{PARKSETC}}$ | 0.0078 (0.0126) | 0.0063 (0.0095) | 0.0039 (0.0096) |
| $b_{\mathrm{WOODS}}$ | 0.0448*** (0.0093) | 0.0062 (0.0093) | 0.0069 (0.0096) |
| $b_{\mathrm{MEETING}}$ | −0.0660*** (0.0257) | −0.0229 (0.0232) | −0.0276 (0.0229) |
| $b_{\mathrm{BARSETC}}$ | 0.0000 (0.0005) | 0.0001 (0.0004) | 0.0001 (0.0004) |
| $b_{\mathrm{RESIDENTIAL}}$ | 0.0678** (0.0304) | 0.0250 (0.0276) | 0.0339 (0.0276) |
| $b_{\mathrm{INDUSTRIAL}}$ | −0.0113*** (0.0023) | −0.0040** (0.0020) | −0.0041** (0.0020) |
| $b_{\mathrm{UNDERUSED}}$ | −0.0843*** (0.0215) | −0.0466** (0.0207) | −0.0401** (0.0206) |
| $b_{\mathrm{PAVED}}$ | 0.4768*** (0.1722) | 0.1238 (0.1634) | 0.2098* (0.1619) |
| $b_{\mathrm{TIME}}$ | −0.2600*** (0.0346) | −0.1190*** (0.0395) | −0.1475*** (0.0386) |
| $b_{\mathrm{STORES}}$ | −0.0287 (0.0289) | −0.0285 (0.0223) | −0.0210 (0.0235) |
| $b_{\mathrm{RESTAURANTS}}$ | 0.0317 (0.0301) | 0.0116 (0.0241) | 0.0042 (0.0242) |
| $c_1$ | | 0.9325*** (0.0306) | 0.8868*** (0.0392) |
| $c_2$ | | 12.4134*** (0.5840) | 0.5498*** (0.0757) |
| $L(\mathbf{c})$ | −2321.4303 | −2276.2982 | −2283.2320 |

注）***，**，*は，漸近的な$t$検定において片側1，5，10%の水準で有意であることを表わしている．
（ ）内の数値は標準誤差である．

**表4　不動産価格の予測**

| 実 現 値 | 予 測 値 モデル1 | モデル2 | モデル3 |
|---|---|---|---|
| 493329.19（千種-9） | 661421.94 | 641443.59 | 630758.41 |
| 129501.77（東-2） | 107094.86 | 114621.23 | 110946.56 |
| 111757.34（北-1） | 135703.85 | 139124.05 | 136495.55 |
| 103990.38（西-3） | 89643.07 | 98755.26 | 99996.04 |
| 164479.59（中村-2） | 153275.76 | 160139.02 | 161514.17 |
| 157573.28（昭和-6） | 139896.39 | 158321.86 | 156389.39 |
| 244369.73（瑞穂-7） | 234186.92 | 243453.32 | 244938.68 |
| 65623.38（熱田-1） | 69985.73 | 63557.45 | 64494.75 |
| 180054.16（中川-9） | 152289.01 | 178422.59 | 178942.36 |
| 49501.81（港-6） | 48287.12 | 50929.06 | 51969.97 |
| 94173.20（南-9） | 70767.00 | 84534.17 | 80168.18 |
| 87028.87（守山-4） | 83503.60 | 80999.19 | 79627.40 |
| 72848.81（緑-7） | 80925.47 | 78281.93 | 77998.62 |
| 137027.27（名東-1） | 128965.13 | 138595.07 | 138201.67 |
| 135955.81（天白-5） | 120073.92 | 126625.12 | 124146.35 |

注）（ ）内は「基準地番号」であり，価格の単位は円である．

**表5　特性価格の推定**

| TIME | 推 定 値 モデル1 | モデル2 | モデル3 |
|---|---|---|---|
| 22（千種-9） | −7816.14 | −3470.58 | −4229.24 |
| 23（東-2） | −1210.53 | −593.20 | −711.55 |
| 18（北-1） | −1960.00 | −920.02 | −1118.58 |
| 34（西-3） | −685.45 | −345.74 | −433.84 |
| 27（中村-2） | −1475.86 | −705.99 | −882.41 |
| 30（昭和-6） | −1212.33 | −628.18 | −768.97 |
| 33（瑞穂-7） | −1844.95 | −878.15 | −1094.88 |
| 21（熱田-1） | −866.42 | −360.26 | −453.03 |
| 49（中川-9） | −808.00 | −433.43 | −538.69 |
| 47（港-6） | −267.10 | −128.98 | −163.11 |
| 34（南-9） | −541.11 | −295.95 | −347.81 |
| 30（守山-4） | −723.64 | −321.38 | −391.53 |
| 53（緑-7） | −396.96 | −175.81 | −217.09 |
| 39（名東-1） | −859.69 | −423.01 | −522.72 |
| 47（天白-5） | −664.18 | −320.69 | −389.64 |

注）（ ）内は「基準地番号」であり，TIME と価格の単位は分と円である．

説明変数にかかわる係数パラメターは，変数をサブスクリプトで示して $b$ に付けることによって記されている。自己相関があるとき，それを考慮しないと，$t$ 統計量はインフレートすることがある。モデル 1 では，DIRECTIONS や DOCTORS，WOODS，MEETING や RESIDENTIAL が有意となっている一方，モデル 2 と 3 では，SCHOOLYARDS や INDUSTRIAL，UNDERUSED や PAVED の有意性が低くなっている。これらはインフレーションを例示するものであると言えよう。

ヘドニック価格関数は，不動産の価格を予測したり特性の価格を推定したりする際に利用される。モデ

ルの間で予測値や推定値を比較することは重要になる。推定に用いられなかった不動産を取り上げることは有用であろう。注意すべきことは，一部の不動産を用いないようにすることは，一般化最小2乗推定を行う場合には，受け入れられないかもしれないということである。データは，214組あり，実験において抽出された200組を超えるサイズを有しているものの，不動産価格の予測や特性価格の推定のためにその一部を温存することは難しいと思われた。『地価調査』は，都道府県が基準地を選定して7月1日時点の価格を判定するもので，『地価公示』を補完するために発行されている。そこで，基準地を取り上げてモデルのパフォーマンスを比べることとしよう。愛知県の『地価調査』（昭和60年）と『学区別生活環境調査報告書』を使ってデータを加える。基準地は，「生活環境図」に示されていないので，『名古屋市地価公示・基準地位置図』や愛知県企画部土地利用調整課の『地価調査基準地案内図』（昭和60年10月）に基づいて特定する。第1種住居専用地域や第2種住居専用地域，住居地域に位置し「利用の現況」が「住宅」となっている147の基準地の中から，『地価調査』(昭和59年)と比較して，価格の変動率が各区で最も小さい15の基準地を選び，1㎡当たりの調査価格に地積と割引率を掛けて，昭和60年1月1日時点のPRICEの実現値を補間する[9]。

基準地のPRICEを予測した結果は表4に与えられている。モデル1については(11)式が用いられている一方，モデル2と3については(9)式が用いられている。実現値に対する予測値の平均平方誤差は，モデル1で2.0721D+9になるのに対して，モデル2と3で1.4911D+9と1.3021D+9になる。これは，空間自己相関を考慮しないと，予測の精度が損なわれることを意味する。モデル3に関しては，(6)式と(8)式を使ってPRICEの対数値を予測し，前者による結果を取り上げた。後者による平均平方誤差は1.3023D+9である。平均平方誤差に基づけば，WM4が選好されることとなろう。

表5は，不動産の価格の予測値でヘドニック価格関数の導関数を評価することによって，特性の価格を推定した結果を与えている。特に関心が向けられると思われるTIMEを取り上げることにした。この係数パラメターの推定値は，モデルにかかわりなく，有意で，期待された符号を伴っている。モデル1による特性価格の推定値は，係数パラメターの推定値とPRICEの予測値に違いがあるために，モデル2や3によるものと大きく異なっている。それらの絶対値は，都心までの所要時分が1分短縮されることに対する人々の評価を表わしていて，昭和60年にあっては，空間計量経済学の成果が利用できなかったことを考えれば，ヘドニック価格関数の撹乱項の自己相関は考慮されず，交通関連投資の必要性が過大に評価されたとしても，疑義が唱えられることはなかったということを示唆する。

## 4．結　語

関心の持たれるモデルが線型であるとは限らない。非線型のモデルを対象にして理論分析を試みることが期待される。ヘドニック価格関数においては，線型性が支持されないこともあり得るので，そうした分析の結果に基づいて推定と予測を行うことは意義を有することになろう。その場合，関数の形状を誤って選択することが撹乱項の自己相関に繋がるということにも注意を払う必要がある[10]。

応用分析では，有意な係数パラメターが多くなかったことから，共線性に関して検討を加えなければならないかもしれない。ただし，それはヘドニック価格関数の特定が十分になされた後に考慮されるべき問題であるとすると，非線型性に配慮して撹乱項の自己相関を特定することが先決問題となる[11]。

[9] それらに関して環境指標の値を知ることに困難はなかった．また，変動率は，平均して，1.4%に過ぎなかった．
[10] 加藤（2001, 2005）を参照．
[11] 加藤（2006）を参照．

## 引用文献


Anselin, Luc（2001）. “Spatial Econometrics,” *A Companion to Theoretical Econometrics*, Baltagi, Badi H. ed., Blackwell Publishers, 310–330.

Anselin, Luc and Bera, Anil K.（1998）. “Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics,” *Handbook of Applied Economic Statistics*, Ullah, Aman and Giles, David E. A. eds., Marcel Dekker, 237–289.

Dubin, Robin（2003）. “Robustness of Spatial Autocorrelation Specifications: Some Monte Carlo Evidence,” *Journal of Regional Science*, **43**, 221–248.

Dubin, Robin, Pace, R. Kelley and Thibodeau, Thomas G.（1999）. “Spatial Autoregression Techniques for Real Estate Data,” *Journal of Real Estate Literature*, **7**, 79–95.

加藤尚史（1996）.「ヘドニック価格アプローチによる環境政策の評価」,『日本統計学会誌』, **26**, 287–319.

加藤尚史（2001）.「空間自己相関を考慮したヘドニック価格関数の特定化と推定」,『応用地域学研究』, 6, 99–110.

加藤尚史（2005）.「不動産価格関数の推定と利用について」,『日本統計学会誌』, **34**, 131–161.

Kato, Takafumi（2005）. “A Further Exploration into the Robustness of Spatial Autocorrelation Specifications,” *DEE Discussion Paper Series*, 05–6, Graduate School of Environmental Studies, Nagoya University.

加藤尚史（2006）.「共線性を考慮した不動産価格関数の推定」, mimeo, 名古屋大学大学院環境学研究科.

Kato, Takafumi（2006）. “Estimation and Prediction in Spatial Linear Regression Models,” mimeo, Graduate School of Environmental Studies, Nagoya University.

森杉壽芳（1991）.「費用便益分析」,『公共セクターの効率化』, 金本良嗣・宮島洋編, 東京大学出版会, 71–87.

Pace, R. Kelley, Barry, Ronald and Sirmans, C. F.（1998）. “Spatial Statistics and Real Estate,” *Journal of Real Estate Finance and Economics*, **17**, 5–13.

Palmquist, Raymond B.（1999）. “Hedonic Models,” *Handbook of Environmental and Resource Economics*, van den Bergh, Jeroen C. J. M. ed., Edward Elgar Publishing, 765–776.

Palmquist, Raymond B.（2005）.“Property Value Models,” *Handbook of Environmental Economics*, **2**, Mäler, Karl-Göran and Vincent, Jeffrey R. eds., Elsevier, 763–819.

Sheppard, Stephen（1999）. “Hedonic Analysis of Housing Markets,” *Handbook of Regional and Urban Economics*, **3**, Cheshire, Paul and Mills, Edwin S. eds., Elsevier Science, 1595–1635.