

第2008-09号

共クリギングによる土地取引価格の時空間 内挿に関する研究

東京大学大学院工学系研究科 社会基盤学専攻
講師 井上 亮

平成21年9月

助成研究者紹介

いのうえ りょう

井上 亮

現職：東京大学 大学院工学系研究科 講師（博士(工学)）

主な論文：

- ・清水英範，井上 亮（2008）カルトグラムの作成手法と応用可能性－統計GISの発展と利活用に向けて－．土木計画学研究・論文集．25(1): 1-15.
- ・井上 亮，清水英範，吉田 雄太郎，李 勇鶴（2009）時空間クリギングによる東京23区・全用途地域を対象とした公示地価の分布と変遷の視覚化．GIS－理論と応用．17(1): 13-24.

共同研究者紹介

り よんふ

李 勇鶴

東京大学 大学院工学系研究科 博士課程

しみず えいはん

清水 英範

東京大学 大学院工学系研究科 教授（工学博士）

目次

1. 序論.....	1
1.1 背景.....	1
1.2 本研究の目的.....	2
2. 共クリギング理論.....	3
2.1 最良線形不偏予測量.....	3
2.2 クリギング.....	5
2.3 共クリギング.....	12
2.4 交差検証.....	16
3. 地価情報への共クリギングの適用可能性の検討.....	18
3.1 空間相関を考慮した共クリギングによる地価内挿実証実験.....	18
3.2 時空間相関を考慮した共クリギングによる地価内挿実証実験.....	25
4. 取引価格に関する情報提供.....	30
4.1 利用した地価情報.....	30
4.2 取引地点の推定と鑑定価格の内挿.....	31
4.3 取引価格と鑑定価格の比較例.....	32
4.4 まとめ.....	34
5. 結論.....	38
5.1 本研究の成果.....	38
5.2 今後の課題.....	39
参考文献.....	40

1. 序論

1.1 背景

近年、市場原理によって土地の高度・有効利用を促進する施策の一環として、不動産市場の透明性の向上、特に地価情報の更なる整備と公開の必要性が叫ばれている。市場参加者が関心を持つ土地の価格や動向を知り、他の土地との比較ができなければ、合理的な意思決定を行うことは不可能であるからである。

日本では「地価情報の整備と公開」の役割を、国土交通省による公示地価や都道府県による基準地価など不動産鑑定評価に基づく価格情報（以後、鑑定価格情報）が担ってきた。これらの鑑定価格情報は取引指標となるべく不動産鑑定士の鑑定に基づき公開されてきたが、不動産市場での取引価格との乖離が常に指摘されている。そこで、国土交通省は一定の制限の下に土地の取引価格に関する情報（以後、取引価格情報）の公開方針を決定し、2005年第3四半期から取引価格等に関する調査を実施し、その結果を2006年4月から「土地総合情報システム」上で公開している。

しかし、これら既存の地価情報には、不動産市場の参加者に価格動向情報を提供する上で、必然的な限界が存在する。まず、新たに公開が開始された取引価格情報では、取引価格は取引当事者の売り急ぎや買い急ぎなど個別事情を反映しているため必ずしも取引物件の標準的な価格を表さないこと、また取引当事者の個別事情が情報提供されていないことなどが問題となる。すなわち、現在公開されている取引価格情報だけから不動産市場の動向を把握することは難しい。一方、旧来から提供されてきた鑑定価格情報においても、費用や各種事情により全ての土地の鑑定価格を提示することは不可能で限られた時点・地点の価格情報しか提供できないため、市場参加者が関心を持つ土地の価格やその動向について十分に情報提供することはできない。

すなわち、全ての不動産市場参加者にとって役立つ価格動向を情報提供するためには、既存のアプローチだけでは不十分であり、時空間で蓄積されてきた鑑定価格情報を活用した内挿というプロセスが不可欠である。

また、そもそも不動産市場は、財の同質性や情報の完全性などが成り立たない典型的な不完全競争市場であるため、その動向把握には取引価格情報・鑑定価格情報の両面からの分析が不可欠である。全ての市場参加者が、これら2種類の価格情報の相互比較を通して市場分析を行える環境を整備することは、市場の透明性向上に大きく寄与するものと考えられる。しかし、取引価格情報と鑑定価格情報の比較は容易ではない。

鑑定価格情報は、時間変動を捉えるよう一定の時間間隔で、また、地域を網羅するよう適切な空間間隔で、かつ、近隣を代表する不動産の価格情報を提供するよう標準的な土地を選定し、その価格を継続的に公表している。一方、取引価格情報は、不動産市場における生の取引価格情報の提供を目的としているため、当然、時空間上で偏在している取引地点の情報が提供されている。取引地点の地積や形状などの属性は、必ずしもその近隣を代表する標準的なものとは限らず、鑑定価格情報では情報提供の対象とされていない不整形の土地など悪条件地の価格情報も含まれている。その結果、取引価格情報の時空間上の近

隣には、単純比較が可能な公的地価指標は存在しない可能性が高い。

しかし、ここで前述の、時空間で蓄積された鑑定価格情報の内挿というプロセスを踏むと、取引地点の鑑定価格情報を内挿し、取引価格と比較することが可能になる。すなわち、取引価格情報と鑑定価格情報の相互比較を通して市場分析を簡単に行うことができる環境を提供することが可能になると推察される。

従来、地価の統計学的な予測に関する研究や実務においては、通常最小二乗法を用いた線形回帰モデルを適用するのが一般であった。しかし、通常最小二乗法によって高い精度の予測モデルを作成するには、膨大な量の説明変数とデータ収集が必要であり、時空間内挿にそのまま適用するのは一般に極めて困難である。

一方、空間統計学の分野では、クリギング(Kriging)と呼ばれる空間内挿手法が体系化されてきた。クリギングは、誤差の共分散を距離の関数（共分散関数）を用いて構造化し、これによって任意地点の空間予測を行う手法であり、理論的には最良線形不偏予測量を与えるという極めて優れた特性を有している。さらに近年では、Cressie and Huang(1999)やGeneting(2002)を端緒に、クリギングを時空間予測に拡張した「時空間クリギング」と呼ばれる手法の理論研究が進展している。地価は時間と空間の両方に対し系列相関をもつ傾向があり、時空間クリギングは地価の時空間内挿に有効な手法と考えられる。井上ら(2009)では、東京 23 区内・33 年間の公示地価データに対して時空間クリギングを適用し、空間相関だけを考慮した既存の内挿手法に比べて、非常に高精度の地価内挿が実行可能であることを確認している。

この地価内挿精度を更に向上させる可能性がある方法として、共クリギング(Cokriging)がある。内挿対象の変数（以後、対象変数と記す）に加えて、対象変数と相関を持った補助変数を用いて内挿を行う手法で、自然環境データへの適用は多く見られ、また、空間相関を考慮した地価情報への適用も行われている。前述のように、地価の内挿では、空間相関に加えて時系列相関を考慮することにより、精度を大きく向上させることが期待される。

1.2 研究の目的

本研究では、共クリギングを時空間相関を考慮するモデルに拡張し、地価内挿精度の向上が可能かどうか実証的に検討を行う。

また、本研究の最終目標は、地価内挿を通して、鑑定価格情報・取引価格情報の双方から全ての不動産市場参加者が市場分析可能な環境を整備し、不動産市場の透明性の改善に資することにある。しかし、前述のように、取引価格情報は取引当事者の個別事情を大きく反映しているため、単独で公開しても、必ずしもデータの閲覧者にとって有意義な情報を提供できるとは限らない。そこで、本研究で提案する地価内挿手法を活用し、取引価格情報と鑑定価格情報の内挿値を比較した情報を提供することを提案し、その情報提供イメージを示す。

2. 共クリギング理論

本章では、空間統計学で提案されてきた内挿手法であるクリギング、および、その一種である共クリギングについて、その理論を解説する。なお、本研究では、線形回帰モデルの誤差の共分散を共分散関数で構造化して内挿を行う普遍型クリギング(Universal Kriging)のみを扱う。

まず、2.1 節ではクリギングの基礎となる最良線形不偏予測量(Best Linear Unbiased Predictor: BLUP)について説明する。次に、2.2 節では内挿対象の対象変数とそのモデルを用いて内挿を行うクリギングについて説明し、2.3 節では対象変数とそのモデルに加えて、相関がある補助変数とそのモデルを用いて内挿を行う共クリギングについて、クリギングと比較を通してその相違点を中心に説明する。また、2.4 節で空間クリギングの時空間への拡張手法について説明する。最後に、2.5 節ではクリギングによる地価情報内挿に関する既往研究を簡単に纏める。

2.1 最良線形不偏予測量

最良線形不偏予測量は、予測量は観測値の線形結合（線形性）、予測量の期待値と真の値が一致（不偏性）、予測誤差分散が最小（最良性）という三つの性質を持つ予測量である。ここでは、Goldberger(1962)に基づき BLUP の導出を行う。

まず、線形回帰モデルとして次のモデルを仮定する。

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V}) \quad (2.1)$$

ただし、 \mathbf{y} は観測値のベクトル、 \mathbf{X} は観測値の説明変数の行列、 $\boldsymbol{\beta}$ はパラメータのベクトル、

$\boldsymbol{\varepsilon}$ は観測誤差のベクトル、 \mathbf{V} は観測誤差の分散共分散行列である。

また、予測モデルとして、以下の仮定をする。

$$\begin{aligned} y_* &= \mathbf{x}'_* \boldsymbol{\beta} + \varepsilon_* \\ \varepsilon_* &\sim N(0, \sigma_*^2) \quad E(\varepsilon_* \boldsymbol{\varepsilon}) = \mathbf{c}_* \end{aligned} \quad (2.2)$$

ただし、 y_* は予測値、 \mathbf{x}_* は予測値の説明変数のベクトル、 ε_* は予測値の攪乱項、 σ_*^2 は予測値の攪乱項の分散、 \mathbf{c}_* は予測値の攪乱項 ε_* と観測誤差ベクトル $\boldsymbol{\varepsilon}$ の共分散ベクトルである。

ここで、 y_* の最良線形不偏予測量 \hat{y}_* の三性質より、以下の式が得られる。

$$\min [\text{var}(y_* - \hat{y}_*)] \quad (\text{最良性}) \quad (2.3)$$

$$\hat{y}_* = \mathbf{a}'\mathbf{y} \quad (\text{線形性}) \quad (2.4)$$

$$E(y_* - \hat{y}_*) = 0 \quad (\text{不偏性}) \quad (2.5)$$

ただし、 \mathbf{a} は重みベクトル

まず、式(2.1)(2.4)(2.5)より

$$\begin{aligned} E(y_* - \hat{y}_*) &= E(\mathbf{x}'_*\boldsymbol{\beta} + \varepsilon_* - \mathbf{a}'\mathbf{X}\boldsymbol{\beta} - \mathbf{a}'\boldsymbol{\varepsilon}) \\ &= (\mathbf{a}'\mathbf{X} - \mathbf{x}'_*)\boldsymbol{\beta} \\ &= 0 \end{aligned} \quad (2.6)$$

となる。式(2.6)が**0**にかかわらず常に成り立つためには

$$\mathbf{a}'\mathbf{X} - \mathbf{x}'_* = 0 \quad (2.7)$$

でなければならない。この条件のもとに予測誤差の分散

$$\text{var}(y_* - \hat{y}_*) = \mathbf{a}'\mathbf{V}\mathbf{a} - 2\mathbf{a}'\mathbf{c}_* + \sigma_*^2 \quad (2.8)$$

が最小となるための必要十分条件は、式(2.7)(2.8)より導かれる

$$f(\mathbf{a}, \lambda) = \mathbf{a}'\mathbf{V}\mathbf{a} - 2\mathbf{a}'\mathbf{c}_* + \sigma_*^2 + 2(\mathbf{a}'\mathbf{X} - \mathbf{x}'_*)\lambda \quad (2.9)$$

を **\mathbf{a} および λ** について偏微分した式(2.10)が共に0であることである。

$$\begin{aligned} \frac{\partial f(\mathbf{a}, \lambda)}{\partial \mathbf{a}} &= 2\mathbf{V}\mathbf{a} - 2\mathbf{c}_* + 2\mathbf{X}\lambda \\ \frac{\partial f(\mathbf{a}, \lambda)}{\partial \lambda} &= 2(\mathbf{X}'\mathbf{a} - \mathbf{x}_*) \end{aligned} \quad (2.10)$$

ただし、 λ はラグランジェの未定乗数である。式(2.10)を行列で表すと、式(2.11)になる。

$$\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{c}_* \\ x_* \end{bmatrix} \quad (2.11)$$

式(2.11)を解くと、

$$\hat{\mathbf{a}}' = \mathbf{c}'_*[\mathbf{I} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}']\mathbf{V}^{-1} + \mathbf{x}'_*(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1} \quad (2.12)$$

が求められる。式(2.12)を式(2.4)に代入すると

$$\begin{aligned} \hat{y}_* &= \hat{\mathbf{a}}'\mathbf{y} \\ &= \mathbf{c}'_*[\mathbf{I} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}']\mathbf{V}^{-1}\mathbf{y} + \mathbf{x}'_*(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \end{aligned} \quad (2.13)$$

となり、一般化最小2乗推定量 $\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ で書き換えると、最良線形不偏予測量は式(2.14)となる。

$$\hat{y}_* = \mathbf{x}'_* \hat{\boldsymbol{\beta}}_{\text{GLS}} + \mathbf{c}'_* \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}_{\text{GLS}}) \quad (2.14)$$

このように、予測量は観測値の線形結合（線形性）、予測量の期待値と真の値が一致（不偏性）、予測誤差分散が最小（最良性）という三つの性質から、式(2.14)の最良線形普遍予測量が導かれた。ただし、 \mathbf{c} 、 \mathbf{V} が既知であることが前提となっている。

なお、分散共分散行列 \mathbf{V} は半正定値符合行列(Semi Positive Definite Matrix)となる。例えば、確率変数ベクトル \mathbf{u} に対して、任意のベクトル \mathbf{a} による線形結合 $\mathbf{a}'\mathbf{u}$ を計算する。このスカラー $\mathbf{a}'\mathbf{u}$ の分散 $\text{var}(\mathbf{a}'\mathbf{u})$ は分散の定義から0以上の値になるが、確率変数ベクトルの分散共分散行列 $\text{var}(\mathbf{u})$ は任意のベクトル \mathbf{a} に対して、式(2.15)の関係を持つ。

$$\begin{aligned} \text{var}(\mathbf{a}'\mathbf{u}) &= E \left[(\mathbf{a}'\mathbf{u} - E(\mathbf{a}'\mathbf{u})) (\mathbf{a}'\mathbf{u} - E(\mathbf{a}'\mathbf{u}))' \right] \\ &= \mathbf{a}' E \left[(\mathbf{u} - E(\mathbf{u})) (\mathbf{u} - E(\mathbf{u}))' \right] \mathbf{a} \\ &= \mathbf{a}' \text{var}(\mathbf{u}) \mathbf{a} \\ &\geq 0 \end{aligned} \quad (2.15)$$

すなわち、確率変数ベクトルの分散共分散行列は半正定値符合行列となる、

通常、誤差の相関構造は不明で、何らかのモデルを仮定し、共分散ベクトル \mathbf{c} 、分散共分散行列 \mathbf{V} を構造化しなければならない。クリギングは誤差の共分散を距離の関数で表すことにより、これらを構造化する手法である。次節に、クリギングについて記す。

2.2 クリギング

クリギングは、空間統計学の分野で議論されてきた空間内挿手法の一つで、元々は鉱床の空間分布の予測（内挿）に使われてきた。現在では、地質分野にとまらず、環境科学、水文学、森林学、水産学など自然環境データを扱う学問領域で広く応用されているほか、社会経済データへの応用も見られるようになってきている。

クリギングでは、距離が近い点のデータは大きな類似性を持つという空間相関を、変数間の共分散を距離の関数として表現することで構造化する。共分散関数のパラメータ推定を通して、式(2.14)の共分散ベクトル \mathbf{c} 、分散共分散行列 \mathbf{V} を推定し、任意の地点に対して経験最良線形不偏予測量(Empirical (or Estimated) Best Linear Unbiased Predictor: EBLUP)を求める手法である。クリギングでは対象変数の自己相関を利用して内挿（予測）を行うが、共クリギ

ングでは、対象変数と相関を持つ他の補助変数も同時に利用して内挿を行う。本節ではクリギングについて記し、次節では共クリギングについてクリギングとの対比を通して説明する。

なお、クリギングに関する理論に関しては、Cressie (1993), Goovaerts (1997), Chiles and Delfiner (1999), Webster and Oliver (2001), 間瀬・武田 (2001), Wackernagel (2003)が詳しい。

(1) 確率場と定常性

クリギングでは線形回帰モデルの残差 (residual) を確率場 (random function) の実現値と見なす。すなわち、領域 $D \subset \mathbb{R}^d$ 上の確率場 $Z = \{Z(\mathbf{s}); \mathbf{s} \in D\}$ を考えると、残差は観測位置

$\mathbf{s}_1, \dots, \mathbf{s}_n$ における確率変数 $Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n)$ の実現値 $z(\mathbf{s}_1), \dots, z(\mathbf{s}_n)$ である。確率場の概念は、

図 2-1 に示したように領域化(regionalization)とランダム性(randomness)の二つの観点から生まれる(Wackernagel, 1998)。

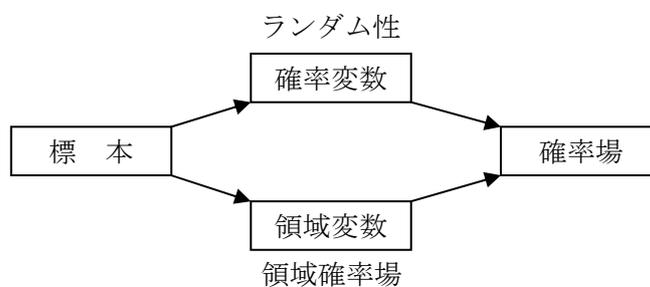


図 2-1 確率場モデル

理論分析のために、確率場には定常性の仮定が設定される。定常性は与えられた n 個の点の組が領域のある部分から他の部分に移動しても、確率場の特徴が不変であることを意味する。このことは、移動不変と呼ばれる。

任意の n 点の組 $\mathbf{s}_1, \dots, \mathbf{s}_n$ および任意のベクトル \mathbf{h} に対しても次式が成立するとき、確率場は強定常(strictly stationary)であるといわれる。すなわち、任意の方向に点は位置を移動しても多次元分布に変化がない場合をいう。

$$F_{\mathbf{s}_1, \dots, \mathbf{s}_n}(z_1, \dots, z_n) = F_{\mathbf{s}_1 + \mathbf{h}, \dots, \mathbf{s}_n + \mathbf{h}}(z_1, \dots, z_n) \quad (2.16)$$

しかし、強定常性はかなり厳しい性質であるため、実際の応用では比較的緩和された2次定常性(second-order stationarity)と固有定常性(intrinsic stationarity)の仮定を設定する。強定常性と比べて、2次定常性と固有定常性では、領域において点の対 $\{\mathbf{s}_1, \mathbf{s}_2\}$ のみを考えるととみに、その完全な分布ではなく1次および2次モーメントのみを考慮する。具体的に、2次定常性は変数に対して、固有定常性は対をなす2点間の差分値に対して1次および2次モーメントの定常性を仮定する。

(2) 共分散関数とバリオグラム

確率場 Z が 2 次定常性をもつとは、任意の $\mathbf{s}, \mathbf{s}_1, \mathbf{s}_2 \in D$ で次の関係が成り立つことである。

$$E\{Z(\mathbf{s})\} = \mu \quad (2.17)$$

$$\text{cov}\{Z(\mathbf{s}_1), Z(\mathbf{s}_2)\} = C(\mathbf{s}_1 - \mathbf{s}_2) \quad (2.18)$$

相対位置 $\mathbf{s}_1 - \mathbf{s}_2$ のみに依存する関数 C を共分散関数またはコバリオグラム(covariogram)とよぶ。当然、分散 $\text{var}\{Z(\mathbf{s})\}$ は \mathbf{s} によらず、一定値 $C(\mathbf{0})$ になる。また、共分散関数は偶関数であり、半正定値符号性をもつ。その証明としては、任意の点の組 \mathbf{s}_α および任意の重み係数の組 w_α に対して必然的に次式が成り立つからである。

$$\text{var}\left(\sum_{\alpha=0}^n w_\alpha Z(\mathbf{s}_\alpha)\right) = \sum_{\alpha=0}^n \sum_{\beta=0}^n w_\alpha w_\beta C(\mathbf{s}_\alpha - \mathbf{s}_\beta) = \mathbf{w}'\mathbf{C}\mathbf{w} \geq 0 \quad (2.19)$$

また、確率場 Z が固有定常性をもつとは、次の性質を満たすことである。すなわち、任意の $\mathbf{s}_1, \mathbf{s}_2 \in D$ に対して、

$$E\{Z(\mathbf{s}_1) - Z(\mathbf{s}_2)\} = 0 \quad (2.20)$$

$$\text{var}\{Z(\mathbf{s}_1) - Z(\mathbf{s}_2)\} = 2\gamma(\mathbf{s}_1 - \mathbf{s}_2) \quad (2.21)$$

が成り立つ。

相対位置のみに依存する 2γ をバリオグラム(variogram)、また γ をセミバリオグラム(semivariogram)とよぶ。セミバリオグラムは、偶関数かつ $\gamma(\mathbf{0}) = 0$ であり、式(2.22)の通り、条件付負定値符号関数である。

$$\text{var}\left(\sum_{\alpha=0}^n w_\alpha Z(\mathbf{s}_\alpha)\right) = -\sum_{\alpha=0}^n \sum_{\beta=0}^n w_\alpha w_\beta \gamma(\mathbf{s}_\alpha - \mathbf{s}_\beta) \geq 0, \left(\sum_{\alpha=0}^n w_\alpha = 0 \text{ の場合}\right) \quad (2.22)$$

また、2次定常性をもつ確率場は固有定常性ももつ。

$$\gamma(\mathbf{h}) = \frac{1}{2} \text{var}\{Z(\mathbf{s} + \mathbf{h}) - Z(\mathbf{s})\} = C(\mathbf{0}) - C(\mathbf{h}) \quad (2.23)$$

$$C(\mathbf{h}) = \gamma(\infty) - \gamma(\mathbf{h}) \quad (2.24)$$

$Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n)$ が与えられたとき、共分散関数とセミバリオグラムはそれぞれ式(2.25), 式(2.26)を用いて推定する.

$$\hat{C}(\mathbf{h}) = \frac{1}{|N(\mathbf{h})|} \sum_{N(\mathbf{h})} \{Z(\mathbf{s}_i) - \bar{Z}\} \{Z(\mathbf{s}_j) - \bar{Z}\} \quad (2.25)$$

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2|N(\mathbf{h})|} \sum_{N(\mathbf{h})} (Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2 \quad (2.26)$$

ただし, $\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z(\mathbf{s}_i)$, $N(\mathbf{h}) := \{(i, j); \mathbf{s}_i - \mathbf{s}_j = \mathbf{h}\}$, $|N(\mathbf{h})|$ は対の総数

クリギングによる内挿では, 分散共分散行列などを計算するために共分散関数を求める必要がある. しかし, 共分散関数の推定量(式(2.25))は次の欠点を持つ.

- (1) $\hat{C}(\mathbf{h})$ は不偏ではないが, $\hat{\gamma}(\mathbf{h})$ は不偏である.
- (2) $\hat{C}(\mathbf{h})$ は本来不要な推定値 \bar{Z} を含み, その分共分散の推定精度は悪くなる.

そのため, 通常, 式(2.26)からセミバリオグラムを推定し, 共分散関数を推定する.

(3) セミバリオグラムのモデル化

クリギングを適用する場合, 等方性(isotropy)を仮定することが多い. 等方性を仮定すると, セミバリオグラムは距離のみの関数で表現できる(式(2.27)). なお, 異方性(anisotropy)をもつセミバリオグラムのモデル化については詳述しないが, 座標にアフィン変換をかけて等方性を仮定できる座標軸に変換することなどが行われる.

$$\gamma(\mathbf{h}) = \gamma^0(\|\mathbf{h}\|) \quad (2.27)$$

γ^0 は図3-2-2に示したような形状をもつことが多い. セミバリオグラムの原点におけるジャンプ量をナゲット効果(nugget effect)とよぶ. ナゲット効果は測定誤差やデータから把握不可能な微視規模変動(microscale variantion)などに起因するといわれる. また, セミバリオグラムの値が一定値以上の距離 h で定数になるとき, その値をシル(sill), その距離 h をレンジ(range)とよぶ. シルの存在は確率場が2次定常性をもつことを示唆し, セミバリオグラムと対応する共分散関数が存在することを意味する. シルとナゲット効果の比が大きいほど, またレンジが大きいほど, 確率場は強い空間従属性をもつ.

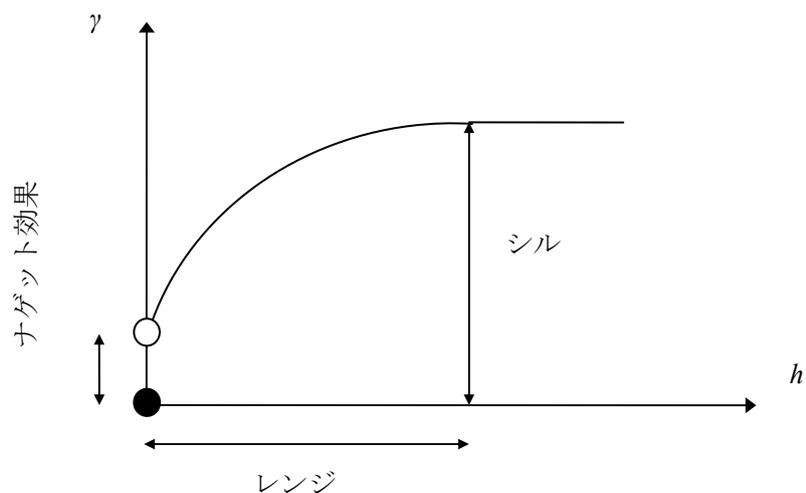


図 2-2 典型的なセミバリオグラム

以下に、各種のモデル化に使用されるセミバリオグラムの例を示す。モデルパラメータ $\theta = [\theta_0, \theta_1, \theta_2]$ および次元 d に対して、指定範囲内で条件付き負定値符号性が成り立つ。

- 1) ナゲット効果モデル (nugget effect model) ($\theta_0 \geq 0, d \geq 1$)

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0, & \|\mathbf{h}\| > 0 \\ 0, & \|\mathbf{h}\| = 0 \end{cases}$$

- 2) 線形モデル (linear model) ($\theta_0, \theta_1 \geq 0, d \geq 1$)

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0 + \theta_1 \|\mathbf{h}\|, & \|\mathbf{h}\| > 0 \\ 0, & \|\mathbf{h}\| = 0 \end{cases}$$

- 3) ベキ乗モデル (power model) ($\theta_0, \theta_1 \geq 0, 0 < \theta_2 < 2, d \geq 1$)

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0 + \theta_1 \|\mathbf{h}\|^{\theta_2}, & \|\mathbf{h}\| > 0 \\ 0, & \|\mathbf{h}\| = 0 \end{cases}$$

- 4) 球形モデル (spherical model) ($\theta_0, \theta_1 \geq 0, \theta_2 > 0, 1 \leq d \leq 3$)

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0 + \theta_1[1.5\|\mathbf{h}\|/\theta_2 - 0.5(\|\mathbf{h}\|/\theta_2)^3], & 0 < \|\mathbf{h}\| \leq \theta_2 \\ \theta_0 + \theta_1, & \|\mathbf{h}\| > \theta_2 \\ 0, & \|\mathbf{h}\| = 0 \end{cases}$$

5) 指数モデル (exponential model) ($\theta_0, \theta_1 \geq 0, \theta_2 > 0, d \geq 1$)

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0 + \theta_1[1 - \exp(-\|\mathbf{h}\|/\theta_2)], & \|\mathbf{h}\| > 0 \\ 0, & \|\mathbf{h}\| = 0 \end{cases}$$

6) ガウスモデル (Gaussian model) ($\theta_0, \theta_1 \geq 0, \theta_2 > 0, d \geq 1$)

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0 + \theta_1[1 - \exp(-\|\mathbf{h}\|/\theta_2)^2], & \|\mathbf{h}\| > 0 \\ 0, & \|\mathbf{h}\| = 0 \end{cases}$$

以上のセミバリオグラムモデルの中で、線形とべき乗モデルにはシルが存在しないため、対応する共分散関数も存在しない。また、それ以外のモデルではシルが $\theta_0 + \theta_1$ になる。但し、指数型とガウス型モデルでは、シルに対応するレンジが存在しないため、セミバリオグラムの値がシルの 95% になるときの距離を実用的レンジ (practical range) とする。例えば、指数型モデルの実用的レンジは $3\theta_2$ にする。

(4) セミバリオグラムのパラメータの推定

セミバリオグラムのパラメータの推定に関しては様々な手法が提案されているが、こちらでは代表的な手法である経験セミバリオグラムに基づいた推定手法について説明する。

➤ 経験セミバリオグラム (empirical semivariogram)

まず、データを用いて、経験セミバリオグラムを計算する (式(2.28))。但し、同じ距離離れた点の組合せはほとんど存在しないので、対応するデータの対 $N(h)$ が少なく、分散の推定が不安定になる。そのため、等方性を仮定した上、距離 h を K 個の区間 $I_1 = (0, R_1], I_2 = (R_1, R_2], \dots, I_K = (R_{K-1}, R_K]$ に分け、区間毎の経験セミバリオグラムを次式で計算する。

$$\hat{\gamma}^0(h_k) = \frac{1}{2|N_k|} \sum_{N_k} \{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\}^2 \quad (2.28)$$

ここで、 $N_{\mathbf{x}} = \{(i, j) \mid \|\mathbf{s}_i - \mathbf{s}_j\| \in I_{\mathbf{x}}\}$ となる (i, j) の対の集合である。区間の幅は、小さすぎると推定が不安定になり、逆に大きすぎると空間従属性が把握できない。それで、 $I_{\mathbf{x}}$ の決定の際には、以下の目安がある (Journel and Højbrechts, 1978)。

- (1) それぞれの区間 $I_{\mathbf{x}}$ において、 $N_{\mathbf{x}}$ の数が 30 より多い。
- (2) パラメータ推定において、 r_{ij} が $\max(r_{ij})$ の 1/2 以上になる (i, j) の対は用いない。

➤ 経験セミバリオグラムに基づいた理論セミバリオグラムのパラメータ推定

理論セミバリオグラムのパラメータ θ は経験セミバリオグラムに適合して推定する (図 2-3)。この推定手法として、式(2.29)の最小二乗基準が提案されている。

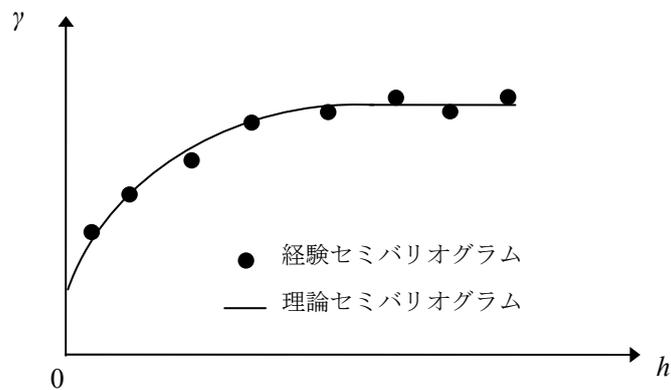


図 2-3 経験セミバリオグラムへの理論セミバリオグラムの当てはめ

$$RSS(\theta) = \sum_{k=1}^K \{\hat{\gamma}^0(h_k) - \hat{\gamma}^0(h_k; \theta)\}^2 \quad (2.29)$$

しかし、式(2.29)は経験セミバリオグラム $\hat{\gamma}^0(h_k)$ の不均一分散を考慮できないことから、Cressie (1985)より式(2.30)の重み付き最小二乗基準が提案されている。内挿を行う際は、セミバリオグラムの原点付近の挙動がきわめて重要であるため、重み付き最小二乗基準はその原点付近のあてはまりを重視しているといえる。

$$WRSS(\theta) = \sum_{k=1}^K \frac{|N_k|}{\hat{\gamma}^0(h_k; \theta)^2} \{\hat{\gamma}^0(h_k) - \hat{\gamma}^0(h_k; \theta)\}^2 \quad (2.30)$$

➤ 理論セミバリオグラムのパラメータ推定手順

上記の手法で理論セミバリオグラムのパラメータ推定を行うには、その基となる経験セミバリオグラムを計算が必要で、線形回帰モデルの残差を求める必要がある。残差の計算には回帰モデルのパラメータの推定値が必要だが、線形回帰モデルのパラメータの一般化最小二

乗推定量となる $\hat{\beta}_{GLS} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ から分かるように、推定には理論セミバリオグラムから求める分散共分散行列が必要である。

本研究では、反復計算により理論セミバリオグラムのパラメータを推定する手法を採用する（図2-4）。

まず、線形回帰モデルのパラメータをOLSによって推定する。その推定結果を用いて線形回帰モデルの残差を算出し、経験セミバリオグラムを作成する。次に、経験セミバリオグラムから、重み付き最小二乗基準により理論セミバリオグラムのパラメータを推定し、共分散関数を設定する。共分散関数を用いて分散共分散行列 \mathbf{V} を定め、一般化最小二乗法で線形回帰モデルのパラメータを推定する。

さらに、その線形回帰モデルのパラメータを用いて残差を算出し、経験セミバリオグラムを算出する。以上の手順を線形回帰モデルのパラメータがある程度収束するまで繰り返し計算を行うことで、理論セミバリオグラムと線形回帰モデルのパラメータを同時推定できる。

以上の内容より、理論セミバリオグラムを求めることができ、クリギングによる経験最良線形不偏予測を行うことができる。

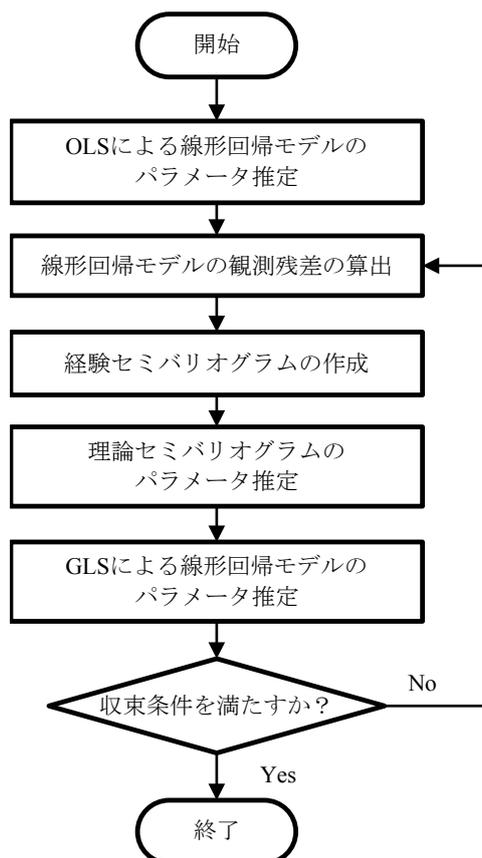


図 2-4 理論セミバリオグラムのパラメータ推定手順

2.3 共クリギング

前節では、対象変数を用いて内挿を行うクリギングについて説明したが、本節では対象変数に加えて、対象変数と相関のある補助変数も用いて内挿を行う共クリギングについて説明する。

共クリギングは、クリギングの多変数への拡張に過ぎず、その理論は共通部分が多い。前節の内容を基として、共クリギングの理論を説明する。また、説明の利便性を考慮して、補助変数が一つの場合を例として説明する。

(1) 最良線形不偏予測量

2.1節では、一変数の場合を例に最良線形不偏予測量について説明したが、二つの変数（一つの対象変数と一つの補助変数）の場合は、以下のように記述される。

まず、線形回帰モデルの仮定は以下ようになる。

$$\begin{aligned} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} \\ E \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} &= \mathbf{0}, \quad E \left(\begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_1 & \boldsymbol{\varepsilon}_2 \end{bmatrix} \right) = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \end{aligned} \quad (2.21)$$

ここで、 \mathbf{y}_i は観測値のベクトル、 \mathbf{X}_i は観測値の説明変数の行列、 $\boldsymbol{\beta}_i$ はパラメータのベクトル、 $\boldsymbol{\varepsilon}_i$ は観測誤差のベクトルであり、 $i=1$ は対象変数、 $i=2$ は補助変数を指す。また、 \mathbf{V}_{11} と \mathbf{V}_{22} はそれぞれ対象変数と補助変数の観測誤差の自己相関を表す分散共分散行列で、 \mathbf{V}_{12} と \mathbf{V}_{21} は対象変数と補助変数の観測誤差の相互相関を表す相互共分散行列である。

しかし、以下のように書き換えると、式(2.21)は式(2.1)と等しく、最良線形不偏予測に関する議論は共クリギングに対しても応用可能である。

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}'_{12} & \mathbf{V}_{22} \end{bmatrix} \quad (2.22)$$

したがって、最良線形不偏予測量は式(2.33)で表される。

$$\hat{y}_* = \mathbf{x}'_* \hat{\boldsymbol{\beta}}_{\text{GLS}} + \mathbf{c}'_* \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}_{\text{GLS}}) \quad (2.23)$$

但し、 $\mathbf{x}'_* = [\mathbf{x}'_{1*} \quad \mathbf{0}']$, $\mathbf{c}'_* = [\mathbf{c}'_{1*} \quad \mathbf{c}'_{2*}]$,

\mathbf{c}_{1*} : 対象変数の観測誤差と予測値の攪乱項の共分散ベクトル,
 \mathbf{c}_{2*} : 補助変数の観測誤差と予測値の攪乱項の共分散ベクトル,
 $\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{y}$: 分散共分散行列 \mathbf{V} 所与時の推定結果

以上の分析より、共クリギングによる予測では、対象変数と補助変数のそれぞれの分散共分散行列の他にも、両変数間の相互共分散行列を求めなければならない。

(2) 相互共分散と相互バリオグラム

クリギングでは確率場に2次定常性や固有定常性を仮定するのに対して、共クリギングで

は同時2次定常性や同時固有定常性を仮定する。

2変数の確率場における自己・相互共分散関数 $C_{ij}(\mathbf{h})$ と自己・相互バリオグラム $2\gamma_{ij}(\mathbf{h})$ は、それぞれ同時2次定常性と同時固有定常性の仮定の枠組みで、次式のように定義される。

$C_{ij}(\mathbf{h})$ と $\gamma_{ij}(\mathbf{h})$ は、 $i=j$ の時は、それぞれ自己共分散関数(auto covariance)と自己セミバリオグラム(direct semivariogram)と呼び、 $i \neq j$ の時は、それぞれ相互共分散関数(cross covariance)と相互セミバリオグラム(cross semivariogram)と呼ぶ。

$$\begin{cases} E\{Z_i(\mathbf{s})\} = 0 \\ \text{cov}\{Z_i(\mathbf{s}), Z_j(\mathbf{s} + \mathbf{h})\} = C(\mathbf{h}) \end{cases} \quad (2.24)$$

$$\begin{cases} E\{Z_i(\mathbf{s}) - Z_i(\mathbf{s} + \mathbf{h})\} = 0 \\ \text{cov}\{Z_i(\mathbf{s} + \mathbf{h}) - Z_i(\mathbf{s}), Z_j(\mathbf{s} + \mathbf{h}) - Z_j(\mathbf{s})\} = 2\gamma_{ij}(\mathbf{h}) \end{cases} \quad (2.25)$$

$(\mathbf{s}, \mathbf{s} + \mathbf{h}) \in \mathbf{D} : i, j = 1, 2$

一組の自己・相互共分散関数と自己・相互セミバリオグラムは、次式に示すようにそれぞれ正定値符号関数と条件付負定値符号関数となる。

$$\text{var}\left(\sum_{i=1}^2 \sum_{\alpha=0}^n w_{\alpha} Z(\mathbf{s}_{\alpha})\right) = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{\alpha=0}^n \sum_{\beta=0}^n w_{\alpha}^i w_{\beta}^j C_{ij}(\mathbf{s}_{\alpha} - \mathbf{s}_{\beta}) \geq 0 \quad (2.26)$$

$$\text{var}\left(\sum_{i=1}^2 \sum_{\alpha=0}^n w_{\alpha} Z(\mathbf{s}_{\alpha})\right) = -\sum_{i=1}^2 \sum_{j=1}^2 \sum_{\alpha=0}^n \sum_{\beta=0}^n w_{\alpha}^i w_{\beta}^j \gamma_{ij}(\mathbf{s}_{\alpha} - \mathbf{s}_{\beta}) \geq 0 \quad (2.27)$$

$\left(\sum_{\alpha=0}^n w_{\alpha} = 0 \text{ の場合}\right)$

また、同時2次定常性の枠組みで、自己・相互共分散関数と自己・相互セミバリオグラムは次式のような関係を持つ。

$$\gamma_{ij}(\mathbf{h}) = C_{ij}(\mathbf{0}) - C(\mathbf{h}) \quad (2.28)$$

共クリギングでは、クリギングと同じく、まずデータに基づいて式(2.25)の自己・相互セミバリオグラムを推定し、その後、式(2.28)から自己・相互共分散関数を求める。

(3) 自己・相互セミバリオグラムのモデル化

上記のように、一組の自己・相互セミバリオグラムは条件付負定値符号関数であり、自己・相互セミバリオグラムのモデル化の際には、この条件を満たすモデルを作成する必要がある。クリギングの場合は、既に条件付負定値符号性を満たす関数が提案されている。しかし、共クリギングの場合は、一つ一つの自己・相互セミバリオグラムが条件付負定値符号性を満たせばいいのではなく、一組としてこの条件を満たさなければならない。つまり、自己・相互セミバリオグラムはお互いに制約関係を持っているため、モデル化の際にはすべて自己・相互セミバリオグラムを同時に考慮する必要がある。

同方性の仮定の上、自己・相互セミバリオグラムはLMCR (Linear Model of

Co-regionalization)を用いてモデル化されることが多い(Lark, 2006; Bishop and Lark, 2007; Parado-Iguzquiza and Atkinson, 2007). LMCРでは、次式に示すように K 個の基本セミバリオグラム(basic semivariogram)モデルの線形結合で自己・相互セミバリオグラムを表示する. 基本セミバリオグラムモデルとはシルが1且つナケット効果が0 (ナケット効果モデルは除く)となるセミバリオグラムモデルのことである.

$$\begin{bmatrix} \gamma_{11}(h) & \gamma_{12}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) \end{bmatrix} = \begin{bmatrix} b_{11}^0 & b_{12}^0 \\ b_{21}^0 & b_{22}^0 \end{bmatrix} g_0(h) + \dots + \begin{bmatrix} b_{11}^K & b_{12}^K \\ b_{21}^K & b_{22}^K \end{bmatrix} g_K(h) \quad (2.29)$$

ここで、自己・相互セミバリオグラムが条件付負定値符号性を満たす充分条件は、基本セミバリオグラムが条件付負定値符号関数で、且つシル行列 $\begin{bmatrix} b_{11}^0 & b_{12}^0 \\ b_{21}^0 & b_{22}^0 \end{bmatrix}, \dots, \begin{bmatrix} b_{11}^K & b_{12}^K \\ b_{21}^K & b_{22}^K \end{bmatrix}$ がすべて正定値符号行列であることである. また、二つ目の条件は次式と等しい.

$$b_{11}^k \geq 0, b_{22}^k \geq 0 \quad k = 1, \dots, K \quad (2.30)$$

$$b_{11}^k b_{22}^k - b_{12}^k b_{21}^k \geq 0 \rightarrow |b_{12}^k| \leq \sqrt{b_{11}^k b_{22}^k} \quad k = 1, \dots, K \quad (2.31)$$

なお、この条件を基に実用性を考慮して三つのルールが提案されている(Goovaerts, 1997).
Rule1: 相互セミバリオグラムに利用した基本セミバリオグラムモデルは、対象変数と補助変数の自己セミバリオグラムにも利用しなければならない.

$$b_{12}^k \neq 0 \rightarrow b_{11}^k \neq 0 \text{ and } b_{22}^k \neq 0$$

Rule2: ある自己セミバリオグラムに利用されていない基本セミバリオグラムモデルは、相互セミバリオグラムにも利用されではない.

$$b_{11}^k = 0 \text{ or } b_{22}^k = 0 \rightarrow b_{12}^k = 0$$

Rule3: 対象変数と補助変数ともに利用されている基本セミバリオグラムモデルが、必ずしも相互セミバリオグラムにも利用される必要はない.

$$b_{11}^k \neq 0 \text{ and } b_{22}^k \neq 0 \rightarrow b_{12}^k \neq 0 \text{ or } b_{21}^k = 0$$

(4) 自己・相互セミバリオグラムのパラメータ推定

理論自己・相互セミバリオグラムのパラメータ推定に関しては、経験自己・相互セミバリオグラムに基づいた推定手法について説明する.

➤ 経験自己・相互セミバリオグラム

経験自己・相互セミバリオグラムの計算方法は、経験自己セミバリオグラムの計算方法と同じく、同方性を仮定した上距離 h を K 個の区間 $I_1 = (0, R_1], I_2 = (R_1, R_2], \dots, I_K = (R_{K-1}, R_K]$ に分け、区間毎の経験自己・相互セミバリオグラムを次式で計算する.

$$\hat{\gamma}_{ij}(h) = \hat{\gamma}_{ij}^0(h_k) = \frac{1}{2|N_k|} \sum_{N_k} [\{Z_i(\mathbf{s}_l) - Z_i(\mathbf{s}_m)\} \{Z_j(\mathbf{s}_l) - Z_j(\mathbf{s}_m)\}] \quad (2.32)$$

ここで、 $i, j=1, 2$ であり、 N_k は $\|\mathbf{s}_l - \mathbf{s}_m\| \in I_k$ となる (i, j) の対の集合である。 $i = j$ の場合 $\hat{\gamma}_{ij}(h)$

は経験自己セミバリオグラムの推定量となり、 $i \neq j$ の場合 $\hat{\gamma}_{ij}(h)$ は経験相互セミバリオグラムの推定量である。 経験相互セミバリオグラムの推定量から分かるように、 経験相互セミバリオグラムを計算するためには、 同じ地点の両変数のデータ必要となる。

➤ 経験自己・相互セミバリオグラムに基づいた

理論自己・相互セミバリオグラムのパラメータ推定

理論自己・相互セミバリオグラムのパラメータは経験自己・相互セミバリオグラムに当てはめ推定する。 ただし、 求められた理論自己・相互セミバリオグラムは条件付き負定値符号性を満たさなければならない。 この条件を満たす理論自己・相互セミバリオグラムのパラメータ推定手法として、 最も簡単な手法は以下である (Goovaerts, 1997)。

まず、 最小二乗基準や重み付き最小二乗基準などに基づいて対象変数と補助変数の理論自己セミバリオグラムのパラメータを推定する。 すなわち、 LMCR での $\hat{\gamma}_{11}^0$ と $\hat{\gamma}_{22}^0 (k = 1, \dots, K)$ を推定する。

次に、 同じ基準に基づき、 $|\hat{\gamma}_{12}^0| \leq \sqrt{\hat{\gamma}_{11}^0 \hat{\gamma}_{22}^0}$ を満たす $\hat{\gamma}_{12}^0 (k = 1, \dots, K)$ を推定する。 すなわち、

理論相互セミバリオグラムのパラメータを推定する。 この推定手法は、 理論自己セミバリオグラムのパラメータ推定精度を優先に考慮した手法であり、 したがって理論相互セミバリオグラムのパラメータ推定精度はある程度低下する。

そこで、 Goulard and Voltz (1992) は重み付き最小二乗基準に基づいて、 理論自己セミバリオグラムと理論相互セミバリオグラムのパラメータを同時に推定する手法を提案している。 この手法では、 スペクトル分解を利用した繰り返し計算により、 次式のような重み付き最小二乗基準が最少となる理論自己・相互セミバリオグラムのパラメータを推定する。

$$WSS = \sum_{k=1}^K |N_k| \text{tr} \left[\left\{ \mathbf{V} \left(\hat{\mathbf{G}}(h_k) - \mathbf{G}(h_k) \right) \right\}^2 \right] \quad (2.33)$$

ここで、 $\hat{\mathbf{G}}(h_k) = \begin{bmatrix} \hat{\gamma}_{11}^0(h_k) & \hat{\gamma}_{12}^0(h_k) \\ \hat{\gamma}_{21}^0(h_k) & \hat{\gamma}_{22}^0(h_k) \end{bmatrix}$ で k 区間での経験自己・相互セミバリオグラムから

なる行列であり、 $\mathbf{G}(h_k) = \begin{bmatrix} \gamma_{11}(h_k) & \gamma_{12}(h_k) \\ \gamma_{21}(h_k) & \gamma_{22}(h_k) \end{bmatrix}$ で k 区間での理論自己・相互セミバリオグラ

ムからなる行列である。 また、 行列 \mathbf{V} は単位行列を使用する場合が多い。

➤ 理論自己・相互セミバリオグラムのパラメータ推定手順

理論自己セミバリオグラムのパラメータ推定手順と同じく、 理論自己・相互セミバリオ

グラムのパラメータ推定も、反復計算により擬似的に真に近い理論自己・相互セミバリオグラムのパラメータを推定する手法を採用する。その推定手順を図2-5に示す。

まず、理論自己セミバリオグラムのパラメータ推定手順に従って、対象変数と補助変数の線形回帰モデルのパラメータをそれぞれ独立に推定する。その推定結果を用いて対象変数と補助変数のそれぞれの線形回帰モデルの残差を算出し、経験自己・相互セミバリオグラムを作成する。そして、経験自己・相互セミバリオグラムから、重み付き最小二乗基準などにより理論自己・相互セミバリオグラムのパラメータを推定し、関係式を利用して対応する自己・相互共分散関数を計算する。自己・相互共分散関数が求まると、分散共分散行列 \mathbf{V} を定めることができるので、一般化最小二乗推定量により、対象変数と補助変数の線形回帰モデルのパラメータが同時に計算できる。さらに、対象変数と補助変数の線形回帰モデルのパラメータを用いてそれぞれの残差を算出し、経験自己・相互セミバリオグラムを算出する。以上の手順をGLS パラメータがある程度収束するまで繰り返し計算を行うことで、真に近い理論自己・相互セミバリオグラムと対象変数と補助変数の線形回帰モデルのパラメータが推定できる。

以上の内容より、理論自己・相互セミバリオグラムを求めることができ、共クリギングによる経験最良線形不偏予測が可能である。

2.4 交差検証

クリギングによるデータの内挿精度は交差検証(Cross-Validation)により評価されることが多い(Stone, 1974; Geisser, 1975)。交差検証とは、データセットから各標本値を順次取り除き、その位置における値を残りの $n-1$ 個 (n は標本数) の標本を用いて内挿し精度検証を行う方法である(Wackernagel, 2003)。この交差検証を Leave-One-Out Cross Validation と呼ぶ。Leave-One-Out Cross Validation ではすべての標本を一回ずつ内挿するため、 n 回の内挿を行う必要がある。

しかし、標本数が大きい場合は計算量が膨大になるため、この代わりに K -分割交差検証 (K -Fold Cross Validation)を用いることが多い(Zhang, 1993; Bengio, 2004)。 K -分割交差検証では、全部の標本を無作為にかつ均等に K 個のグループに分割し、そのうちの 1 グループを内挿精度の検証用 (選択されたグループの標本位置の値を内挿) とし、残りの $K-1$ 個のグループをモデル構築用とする。それで、内挿を K 回繰り返すことによりすべてのグループが一回ずつ検証用になる。

Leave-One-Out Cross Validationと K -分割交差検証は様々な研究により比較されている(Shao, 1993; Burman, 1989; Zhang, 1993)。その結果、 K -分割交差検証による検証結果は K が大きくなるほど正確になり、 K が標本数 n になる場合最も正確で、Leave-One-Out Cross Validationと同じ水準になる。また、 K -分割交差検証は K が5より大きい場合、Leave-One-Out Cross Validationと同じ水準の検証ができるため、計算量と検証正確性の両方を考慮した場合、 K を5から20間の値に設定するのが効率的である(Zhang, 1993)。

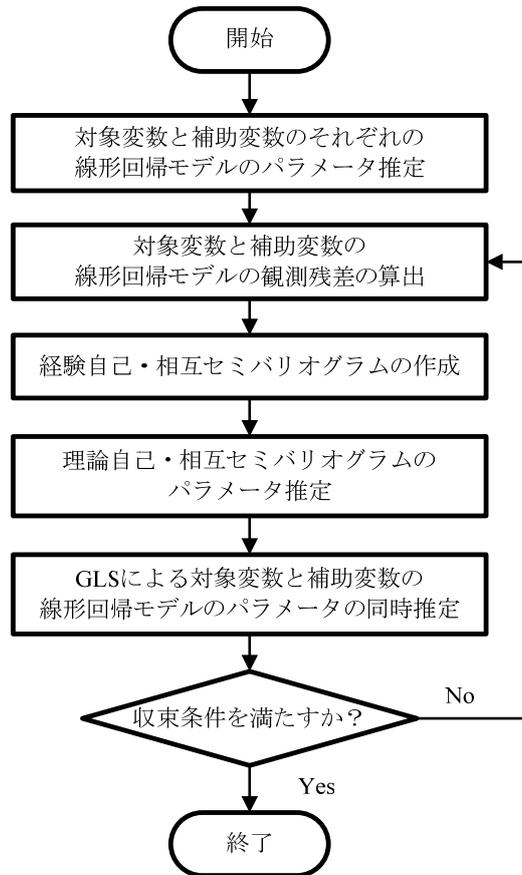


図 2-5 理論自己・相互セミバリオグラムのパラメータ推定手順

3. 地価情報への共クリギングの適用可能性の検討

3.1 空間相関を考慮した共クリギングによる地価内挿実証実験

本節では、時空間共クリギングによる地価内挿の準備段階として、空間共クリギングによる地価内挿の実証実験を行う。

本節の実証実験の主目的は、補助変数の空間分布パターンが内挿精度に与える影響を分析することにある。対象変数と同じ地点の補助変数以外の追加補助変数が一定間隔に分布するグリッド型（図 3-1(a)）と、内挿地点のみに分布するコロケーション型（図 3-1(b)）の二種類の空間共クリギングを用いた地価内挿を行い、その精度を比較・検証する。

本実験では、東京 23 区の 2006 年の住居系用途地域内の公示地価データを対象変数、同年の相続税路線価（住居系）データを補助変数として、両種類の空間共クリギングを用いて地価を内挿し、内挿精度を空間クリギングとの比較を通して評価する。

(1) 実験データ

▶ 公示地価データと路線価データ

本実験では、対象変数として、東京 23 区・2006 年の住居系用途地域内にある公示地価データを用いる。対象の公示地価データは 990 点で、その空間分布と価格水準を図 3-2 に示す。地価は都心から離れるにつれて安くなるが、等距離圏では、南西部で高く、北東部で安い。地価公示点の空間分布は、用途地域の設定により都心部や隅田川周辺の一部の地域では極端に少ないが、それ以外の地域では、ほぼ等間隔で存在する。

また、路線価は道路区間に設定された価格であるが、共クリギングの補助変数として利用するには点データに変換する必要がある。税額の算定時には、ある土地の路線価は、隣接道路の路線価を、その土地の地積、奥行き、地形などの個別要因によって補正する。しかし、本研究では、全ての土地に関して個別要因による補正を行うために必要な情報を用意できないことから、最寄り道路の路線価をその土地の路線価とし、個別要因による価格の補正は行っていない。

なお、本実験で利用した公示地価と路線価の同地点での値の相関係数は、0.999 と非常に強い相関を示している。

▶ 補助変数（路線価データ）の空間分布

共クリギングでは、対象変数と補助変数の相関を構造化する相互バリオグラム関数のパラメータを求めるために、同じ地点に対象変数・補助変数が存在する必要がある。また、共クリギングによる内挿では、主に補助変数が対象変数より密に分布している場合、その内挿精度がクリギングより著しく上昇するとされており (Chica-Olmo, 2007)、対象変数と同地点の補助変数だけを用いた内挿では十分に精度を上げることができない。そこで、本実験では補助変数の分布密度を高めるために、対象変数と同じ地点の補助変数以外にも補助変数を追加する。追加する補助変数は、地価内挿地点に補助変数が分布するコロケーショ

ン型、および、一定間隔で分布するグリッド型の2種類の空間分布を設定し、比較実験を行う。またグリッド型は、100m・200m・300m間隔の補助変数の空間分布を設定した。本論文では以後、補助変数の空間分布により、コロケーション型共クリギング・グリッド型共クリギングと呼ぶ。

なお、グリッド型の追加補助変数データは、東京23区の住居系用途地域内に100m間隔に点を作成し、抽出された地点の最寄りの路線を検索、その路線の路線価を与えた。図3-3に作成した100mグリッド型の追加補助変数の空間分布を示す。また、200mグリッド型と300mグリッド型は100mグリッド型のデータから200m、300m間隔で抽出して作成した。追加補助変数の数は、100mグリッド型では30380点、200mグリッド型では7403点、300mグリッド型では3168点である。

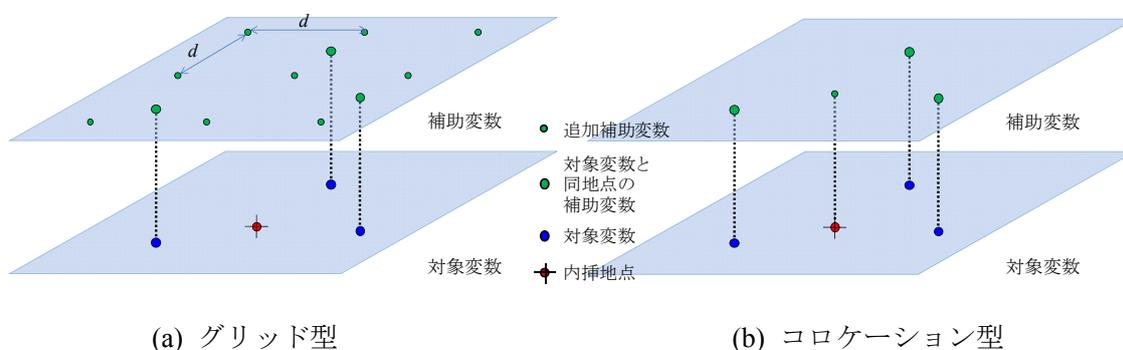


図 3-1 グリッド型とコロケーション型の共クリギング

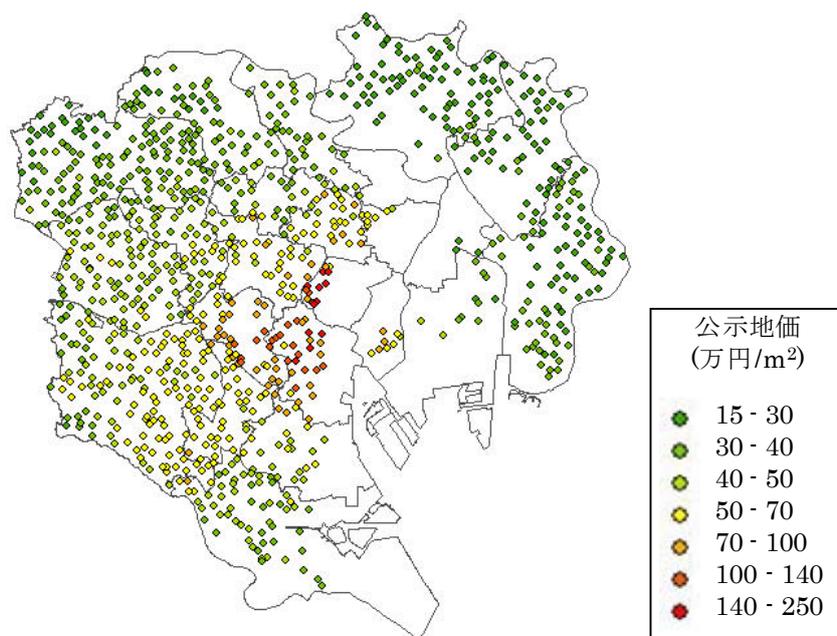


図 3-2 東京 23 区の 2006 年の住居系公示地価の分布



(a) 23 区全体

(b) 本郷周辺拡大図

図 3-3 100m グリッド型の追加補助変数の空間分布 (住居系)

(2) 地価モデルとバリオグラム

➤ 地価モデル

本実験では、対象変数の地価モデルとして両対数型のモデル (式(3.1)) を使用した。

$$\ln(P) = \beta_0 + \sum_{i=1}^5 \beta_i \ln(x_i) + \varepsilon \quad (3.1)$$

但し、 P : 公示地価, x_1 : 主要駅までの鉄道所要時間 (分), x_2 : 最寄り駅までの距離 (m),
 x_3 : 容積率 (%), x_4 : 地積 (m²), x_5 : 前面道路の幅員 (m)

地価公示点の交通条件を表現するために、「最寄り駅」と「最寄り駅迄の距離」の属性を用いている。「最寄り駅」の交通条件を表すため、「最寄り駅」から乗降客数の多い都内主要 5 駅 (新宿・池袋・東京・渋谷・上野) までの鉄道平均所要時間データを作成した。各主要駅までの鉄道所要時間を、平成 17 年実施の第 10 回大都市交通センサスの乗降客数で重み付けし加重平均を行った。各年の鉄道所要時間は、乗り換え案内ソフトウェアである株式会社ヴァル研究所「駅すばあと」を使用し、「不通区間設定機能」で当時の鉄道網を定義した上で現在の時刻表データを用いて最短の乗車時間を算出した。

また、補助変数の地価モデルは、対象変数モデルの説明変数のうち、地積・前面道路幅員を外している。補助変数の説明変数情報は、対象変数 (公示地価) と同じ地点にある補助変数は公示地価の属性情報を利用し、追加補助変数は街区情報、道路網情報、駅情報と容積率情報などに基づいて GIS を用いて算出した結果を利用した。

また、本実験では空間共クリギングと内挿精度を比較するため、空間クリギングによる内挿も行っている。空間クリギングの地価モデルは、式(3.1)の地価モデルを使用した。

➤ バリオグラム

空間共クリギングの理論セミバリオグラムモデルは、自己セミバリオグラム・相互セミバリオグラム共に同じ球形型モデル (式(3.2)) を採用した。また、理論セミバリオグラムモ

デルのレンジは東京 23 区の最長直線距離の約半分である 15km に設定した。比較対象の空間クリギングでも同じ理論セミバリオグラムモデルを使用している。

$$\gamma(\mathbf{h}; \theta) = \begin{cases} \theta_0 + \theta_1[1.5\|\mathbf{h}\|/\theta_2 - 0.5(\|\mathbf{h}\|/\theta_2)^3] & 0 < \|\mathbf{h}\| \leq \theta_2 \\ \theta_0 + \theta_1 & \|\mathbf{h}\| > \theta_2 \\ 0 & \|\mathbf{h}\| = 0 \end{cases} \quad (3.2)$$

(3) 地価モデルとバリオグラムのパラメータ推定結果

前述の地価データと理論バリオグラムモデルを用いて、空間共クリギングによる地価内挿の実験を行い、通常最小二乗法、空間クリギングと比較した。空間共クリギングでは追加補助変数として 100m, 200m, 300m グリッド型とコロケーション型を用いている。

まず、実験結果として、通常最小二乗法(OLS)、空間クリギングと空間共クリギングによる地価モデルのパラメータ推定結果を表 3-1 に示す。空間共クリギングではコロケーション型と 200m グリッド型の結果を示す。三つのグリッド型でのパラメータの推定結果はほぼ等しいため、ここでは 200m グリッド型の結果のみを示している。

まず、通常最小二乗法と空間クリギングのパラメータ推定結果を比較すると、主要駅までの鉄道所要時間、最寄り駅までの距離、地積などのパラメータの絶対値は、空間クリギングで著しく小さい。その原因は、被説明変数に空間相関があるが、空間クリギングでは共分散関数によって説明されるのに対し、通常最小二乗法では同種の空間相関構造を持った説明変数のパラメータを過大に推定することによって説明されるためだと推測される。

次に、空間クリギングと空間共クリギングの対象変数のパラメータを比較する。

表 3-1 地価モデルのパラメータ推定結果

説明変数		OLS		空間クリギング		空間共クリギング (200m グリッド)		空間共クリギング (コロケーション)	
		パラメータ	t 値	パラメータ	t 値	パラメータ	t 値	パラメータ	t 値
対象変数	定数項	15.89	69.4	13.76	60.8	12.48	44.7	13.98	15.4
	主要駅までの 鉄道所要時間	-0.79	-23.5	-0.23	-6.0	-0.04	-1.2	-0.24	-1.5
	最寄り駅までの距離	-0.21	-14.3	-0.08	-10.8	-0.02	-3.1	-0.08	-2.4
	容積率	-0.07	-3.3	-0.06	-4.8	0.03	3.2	0.008	0.2
	前面道路の幅員	0.14	5.7	0.14	15.8	0.04	6.1	0.0004	0.3
	地積	0.23	13.8	0.07	10.7	0.01	2.6	0.005	5.6
補助変数	定数項	---	---	---	---	12.23	51.5	13.75	15.2
	主要駅までの鉄道所 要時間	---	---	---	---	-0.03	-1.7	-0.24	-1.5
	最寄り駅までの距離	---	---	---	---	-0.02	-4.5	-0.08	-2.4
	容積率	---	---	---	---	0.05	11.0	0.01	0.2

注：網かけされたパラメータは 1%水準で有意ではない。

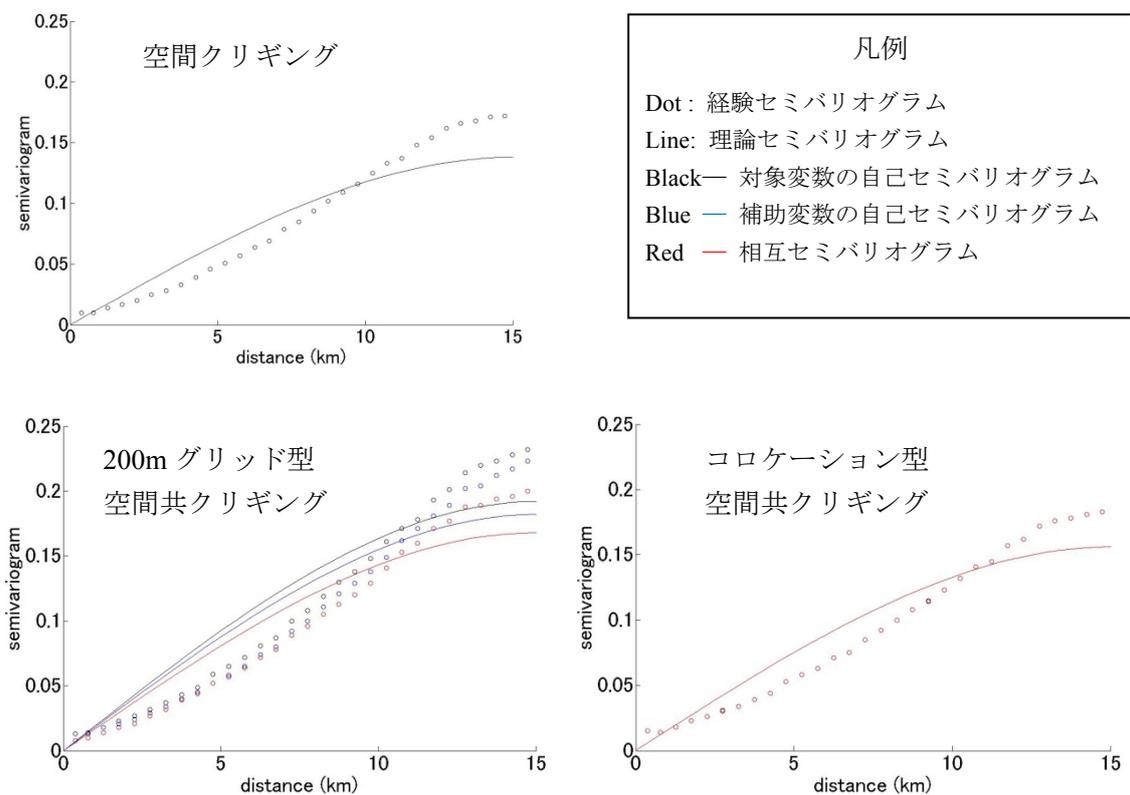


図 3-4 経験セミバリオグラムと理論セミバリオグラム

表 3-2 バリオグラムのパラメータ推定結果

	空間クリギング	200m グリッド型 空間共クリギング			コロケーション型 空間共クリギング		
	自己	自己(対象)	自己(補助)	相互	自己(対象)	自己(補助)	相互
Nugget	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sill	0.138	0.192	0.181	0.168	0.156	0.156	0.156
Range (km)	15	15	15	15	15	15	15

注：自己：自己セミバリオグラム，相互：相互セミバリオグラム，対象：対象変数，補助：補助変数

まず、パラメータの有意性をみると、空間クリギングでは全てのパラメータが1%水準で有意であるのに対して、空間共クリギングでは多くのパラメータが1%水準で有意ではない。

すなわち、空間クリギングでは説明変数によって説明された公示地価の部分が、空間共クリギングでは対象変数と相関が強い補助変数を用いることによって、大半が相互相関によって説明されたことを示している。また、有意であるパラメータの値も、絶対値が空間共クリギングで著しく小さくなっている。その原因も上記と同じく、説明変数によって説明された公示地価の一部が相互バリオグラムによって説明されたためだと考えられる。

次に、空間クリギング，200m グリッド型空間共クリギング，コロケーション型空間共ク

リギングの経験セミバリオグラムと対応する理論セミバリオグラムを図 3-4、理論セミバリオグラムのパラメータの値を表 3-2 に示す。

まず、図 3-4 での空間クリギングと空間共クリギングの経験セミバリオグラムを比較すると、ほぼ同じ形状をしていることが分かる。また、経験セミバリオグラムの原点付近の値（ナゲットと等しい）は原点に近く、遠距離の値（シルと等しい）がナゲットより著しく大きいことから、空間クリギング、空間共クリギングとも空間相関が強いと判断できる。

また、表 3-2 の対象変数の理論自己セミバリオグラムのシルの比較では、空間共クリギングが空間クリギングより若干大きい、空間共クリギングでは地価モデルの誤差の空間相関によって説明される地価の部分が空間クリギングより大きいためであろう。

次に、200m グリッド型空間共クリギングとコロケーション型空間共クリギングを比較すると、三つの経験自己（相互）セミバリオグラムは 200m グリッド型空間共クリギングでは少し異なり、コロケーション型空間共クリギングではほぼ重なっていることが分かる。すなわち、対象変数と補助変数の自己相関構造が非常に類似していることを示す。特に、コロケーション型空間共クリギングでは完全に同じ地点の対象変数と補助変数だけを用いているため、三つの経験自己（相互）セミバリオグラムが重なるということは、両変数の空間相関構造がほぼ一緒であることを示唆している。

(4) 内挿精度の検証

各手法による公示地価の内挿精度を 5-fold cross validation を用いて検証し、表 3-3 に観測値と内挿値の RMSE(Root Mean Square Error)を示す。

まず、300m グリッド型空間共クリギング以外の空間共クリギングは、空間クリギングより高精度の内挿が可能であることが確認された。追加補助変数のグリッド間隔が 300m を超えると、補助変数の効果がほとんど見られない。グリッド型空間共クリギングの精度比較から、追加補助変数の間隔が短くなるほど内挿精度が高くなることが分かり、コロケーション型の空間共クリギングはグリッド型よりも内挿精度が著しく高いことが分かる。

次に、内挿地点から最寄りの補助変数までの距離が、グリッド型空間共クリギングの内挿精度に与える影響を分析するために、グリッド型空間共クリギングの内挿結果を最寄りの補助変数までの距離によって 4 分割し、内挿精度（RMSE）を算出した。100m 型空間共クリギングでは 20m 間隔、200m 型空間共クリギングでは 40m 間隔、300m 型空間共クリギングでは 60m 間隔に分割している。例えば、100m 型空間共クリギングでは 0-20m, 20-40m, 40-60m, 60-80m の 4 組に分け、内挿精度を計算している。図 3-5 に距離別内挿精度を示す。

グリッド型空間共クリギングでは、内挿地点から最寄りの補助変数までの距離が短いほど内挿精度が高い傾向がみられる。また、コロケーション型空間共クリギングは、最寄りの補助変数までの距離がゼロの内挿点と見なすことができる。すなわち、グリッド型空間共クリギングの内挿精度は、追加補助変数データの間隔が短くなればコロケーション型空間共クリギングの RMSE に収束すると推測できる。

表 3-3 各手法による公示地価の内挿精度 (RMSE)

	空間クリギング	300m グリッド型 空間共クリギング	200m グリッド型 空間共クリギング	100m グリッド型 空間共クリギング	コロケーション型 空間共クリギング
RMSE	0.095	0.097	0.085	0.067	0.02

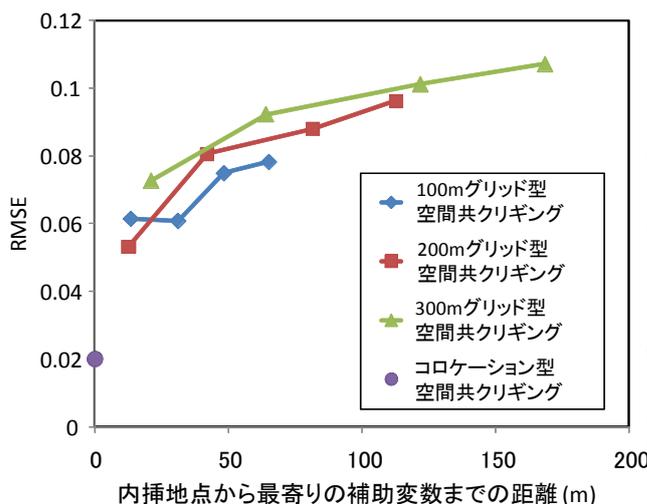


図 3-5 最寄りの補助変数までの距離と内挿精度の関係

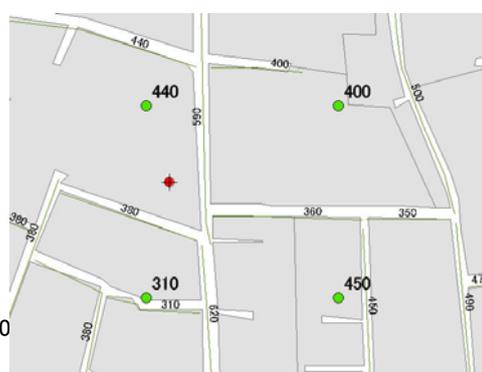


図 3-6 グリッド型空間共クリギングによる地価内挿の問題点

(5) グリッド型とコロケーション型の空間共クリギングの比較

以上の実験結果に基づき、内挿精度、補助変数の数、計算量の三点からグリッド型空間共クリギングとコロケーション型空間共クリギングを比較・分析する。

まず、内挿精度は明らかにコロケーション型空間共クリギングのほうが高い。もちろん、グリッド型空間共クリギングで、追加補助変数の間隔を十分に短くすればコロケーション型空間共クリギングと同水準の内挿ができると考えられる。しかし、グリッド型空間共クリギングには必然的な問題があり、短距離で地価が急激に変化する場合は内挿精度が著しく低下する恐れがある。

例えば、図 3-6 に 100m グリッド型空間共クリギングによる内挿例を示す。赤点は内挿地点、青点は内挿地点周辺の追加補助変数を表す。この例では、内挿地点は広幅員街路に面してその路線価は 590 千円/m²であるが、内挿地点周辺の追加補助変数はより狭い街路に近接しており、その路線価は相対的に低い。したがって、この補助変数に基づいて内挿を行った場合、その内挿値は正しい広幅員街路の路線価を利用した場合より低く内挿される。次に、必要な補助変数の数をモデル構築と内挿の際に分けて分析する。モデル構築の際に、コロケーション型空間共クリギングでは対象変数と同量の補助変数だけを利用するのに対して、グリッド型空間共クリギングでは追加補助変数も加えて利用するため、明らかにコロケーション型空間共クリギングのほうが必要な補助変数の数が少ない。一方、内挿の際

には、コロケーション型空間共クリギングは内挿地点の補助変数を必要とするが、グリッド型空間共クリギングでは必要がないため、コロケーション型空間共クリギングのほうが必要な補助変数の数が多い。

対象変数の数は両手法で同じであるため、計算量は主に補助変数の数によって決まる。当然、補助変数の数が多いグリッド型共クリギングの計算量が、コロケーション型より大きくなる。しかし、内挿の際にグリッド型空間共クリギングではある工夫によって計算量を削減することができる。その方法としては、空間共クリギングでは式(2.23)を用いて予測

値を求めるが、グリッド型空間共クリギングの場合、式(2.23)右辺第二項の $\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\text{GLS}})$

は内挿点の位置に関わらず変化しないため、モデル構築時に一回パラメータ推定を行えば、内挿時にはパラメータ推定結果をそのまま利用できる。しかし、コロケーション型空間共クリギングでは、分散共分散行列 \mathbf{V} に内挿地点の補助変数の影響が入るため、内挿点が変わる度に、再計算が必要である。なお、分散共分散行列 \mathbf{V} の逆行列計算に非常に時間がかかるため、多くの地点の内挿計算をその都度行くと、グリッド型共クリギングによる内挿の方がコロケーション型より計算時間が短くなる。

簡単に纏めると、コロケーション型空間共クリギングは内挿精度、必要な補助変数の数、モデル構築際の計算量などの面で、すべてグリッド型空間共クリギングより優れているが、内挿地点毎に新しくパラメータ推定を行い、分散共分散行列の逆行列を求める必要があるという弱点をもっている。

次節では、時空間共クリギングによる地価内挿への適用可能性について検討する。この実験では、高精度の内挿が可能なコロケーション型を用いる。

3.2 時空間相関を考慮した共クリギングによる地価内挿実証実験

時空間共クリギングの地価内挿への適用可能性を検証するために、1999～2006年の東京23区の住居系公示地価データを対象変数、同期間の路線価データを補助変数として実験を行い、パラメータ推定結果・地価内挿結果を時空間共クリギングと比較する。

(1) 実験データ

対象地域・期間内には住居系の公示地価データは8,010点存在する。また、路線価データは、コロケーション型配置とし、地価公示点最寄りの路線価を使用する。複数の道路と隣接している場合、最寄りの路線価を検索すると本来使用すべき路線価とは異なる値が設定される場合がある。しかし、全地点の路線価を確認することは困難であるため、公示地価と比較して平均的な路線価水準である公示地価の8割より著しく低い路線価だけを抽出して修正した。なお、同一地点の公示地価対数値と修正後の路線価対数値は相関係数が0.999と非常に高い相関がある。

(2) 地価モデルとバリオグラム

公示地価と路線価の地価モデルの説明変数として、主要駅までの鉄道所要時間、最寄り

駅までの道路距離，容積率，地積，前面道路幅員の情報を用いた．日経平均株価は対象期間の1999～2006年では地価との相関が低いため利用していない．地価モデルは，公示地価・路線価共に式(3.2)式を用い，理論自己・相互セミバリオグラム関数は全部同じく式(3.2)を用いた．なお，理論自己・相互セミバリオグラムのレンジは空間上では15km，時間上では4年と設定し，パラメータ推定はLMCRに基づいた重み付き最小二乗基準を用いた．

$$\ln(P) = \beta_0 + \sum_{i=1}^5 \beta_i \ln(x_i) + \varepsilon \quad (3.3)$$

ただし， P : 地価(円/㎡)， x_1 : 主要駅までの鉄道所要時間(分)， x_2 : 最寄り駅までの距離(m)， x_3 : 容積率(%)， x_4 : 前面道路の幅員(m)， x_5 : 地積(㎡)， β_i : パラメータ， ε : 観測誤差

$$\gamma(\mathbf{h}|\theta) = \tau^2 + \sigma_1^2 * Sph(-\|\mathbf{h}\|/\theta_1) + \sigma_2^2 * Sph(-|t|/\theta_2) \quad (3.4)$$

ただし， γ : 理論自己・相互セミバリオグラム， τ^2 : ナゲット， σ_1^2 ， σ_2^2 : 空間・時間のシル， θ_1 ， θ_2 : 空間と時間軸のレンジ， Sph : 球型モデル

(3) 地価モデルとバリオグラムのパラメータ推定結果

まず，時空間クリギングと時空間共クリギングの地価モデルのパラメータ推定結果を，表3-4に示す．両者の比較から，前面道路幅員のパラメータの絶対値だけが時空間共クリギングで時空間クリギングより小さくなったが，他のパラメータの値はほぼ変わらないことが観察できる．ある説明変数のパラメータの絶対値が時空間共クリギングで小さくなるというのは，時空間クリギングではある説明変数によって説明されていた地価形成要因の一部が，時空間共クリギングでは構造化された対象変数と補助変数の地価モデルの誤差の相互相関によって説明されたと解釈することができる．一方，時空間共クリギングの対象変数と補助変数のパラメータは定数項を除いてはほぼ同じであるが，それは公示地価と路線価が強い相関を持っているためだと推察される．

次に，時空間クリギングと時空間共クリギングの理論自己・相互セミバリオグラムのパラメータ推定結果を表3-5に示す．4種類のセミバリオグラムのパラメータはほぼ同じで，時空間相関構造が類似していると考えられる．

ここで，時空間共クリギングの対象変数の自己セミバリオグラムを例として取り上げ，図3-7に示す．空間軸で距離15kmでの経験セミバリオグラムは原点付近の経験セミバリオグラムより著しく大きくなっており，それから地価の空間上の相関は強いと考えられる．一方，時間軸の4年差での経験セミバリオグラムは原点付近の値と比べてほぼ変化がないことから，時系列相関が極めて強く評価されている．この推定されたセミバリオグラムを基にした内挿では，強い時系列相関構造のため前後の年度の影響により値が定まることが予想される．時間の経過と共に影響が減衰する様子が表現されておらず，時系列相関構造を正しく推定しているとは言い難い結果になった．おそらく，データの利用可能性の問題から，対象期間を短く設定せざるを得なかったことが影響しているものと考えられる．この問題に関しては，内挿精度の検証を通して考察する．

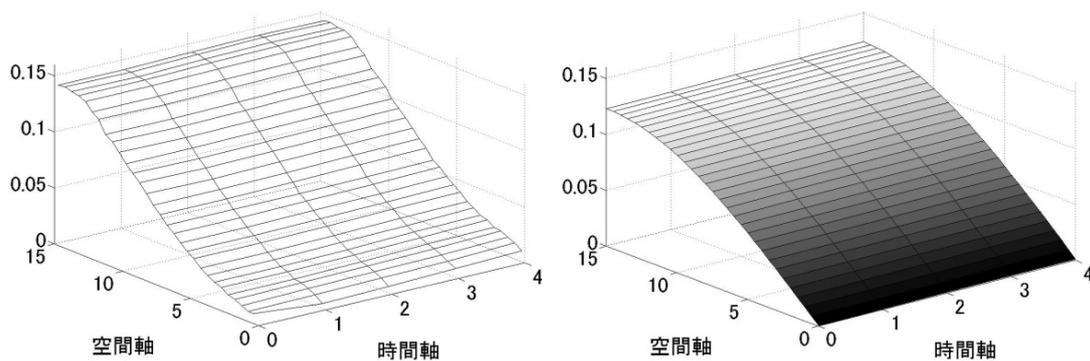
表3-4 地価モデルのパラメータ推定結果

	時空間クリギング (対象変数)	時空間共クリギング (対象変数)	時空間共クリギング (補助変数)
定数項	13.886 (33.7)	13.945(35.1)	13.749 (34.8)
主要駅までの 鉄道所要時間	-0.168 (-5.0)	-0.177 (-6.1)	-0.171 (-6.0)
最寄り駅までの距離	-0.075(-13.2)	-0.074 (-15.3)	-0.074 (-15.4)
容積率	-0.081 (-8.3)	-0.082 (-9.8)	-0.083 (-10.0)
前面道路の幅員	0.081 (5.1)	0.065 (4.2)	0.064 (4.2)
地積	0.047 (4.8)	0.047 (5.1)	0.040 (4.4)

表3-5 理論自己・相互セミバリオグラムのパラメータ推定結果

	時空間クリギング (γ_{11})	時空間共クリギング (γ_{11})	時空間共クリギング (γ_{22})	時空間共クリギング (γ_{12})
ナゲット	0.0001	0.0001	0.0001	0.0001
空間シル	0.115	0.123	0.122	0.122
空間レンジ	15	15	15	15
時間シル	0.001	0.001	0.001	0.001
時間レンジ	4	4	4	4

(注: γ_{11} , γ_{22} : 対象変数と補助変数の自己セミバリオグラム, γ_{12} : 相互セミバリオグラム)



(a) 経験自己セミバリオグラム (b)理論自己セミバリオグラム

図 3-7 対象変数の経験・理論自己セミバリオグラム

(4) 内挿精度の検証

時空間クリギングと時空間共クリギングの内挿精度を5-fold cross validationを用いて比較した。まず、検証の結果、全期間におけるRMSEは時空間クリギングの0.044に対し、時空間

共クリギングでは0.018となっており、時空間共クリギングによる内挿精度が優れていることが分かる。

図3-8に年毎の内挿精度の違いを表現している。なお、比較対象として、各年のデータを用いて、空間相関のみを考慮した空間共クリギングの内挿精度を合わせて表示している。全期間において、時空間共クリギングの内挿精度は、時空間クリギング・空間共クリギングよりもことが示されている。また、時空間クリギングの内挿精度は2002～2003年の対象期間中央部で最も高く、両端（1999年と2006年）では大きく低下する。一方、時空間共クリギングでは全年度でほぼ同水準の内挿精度を保つことが分かる。このように対象期間の両端では、補助変数を利用する時空間共クリギングの効果が著しいことが確認された。

次に、内挿精度の空間分布を確認するため、2006年の内挿残差（内挿価格の対数値-鑑定価格の対数値）を図3-9に示す。

まず、時空間クリギング・時空間共クリギング両方の内挿残差に空間相関が見られることが分かる。本来、クリギングや共クリギングなど時系列相関・空間相関を考慮したモデルで内挿を行うと、その内挿残差には系列相関は残らない。図3-10には、空間クリギングと空間共クリギングによる内挿残差の分布を示すが、空間相関は見られない。このように、相関構造を推定できている時には、内挿残差には系列相関は生じない。

すなわち、今回の実験では、正しく時空間相関構造を推定できていないことが、内挿残差に系列相関が残っている原因だと考えられる。前述のように、この実験ではデータの対象期間が短いという問題があり、特に時系列の相関構造が正しく推定できていない可能性が高い。その結果、相関構造の推定に問題が生じ、内挿残差の相関に影響したと考えられる。しかし、相関構造の推定に問題があったとしても、全体的に見て内挿精度が向上しているということは興味深く、地価内挿時に時空間相関の考慮することが重要であることを示唆していると言えるだろう。

本実験では、データの制約から十分な期間の地価情報を入力できなかったため、時空間相関構造の推定に問題が生じたが、時空間共クリギングにより高精度の内挿が行える可能性について示すことができたと考えている。

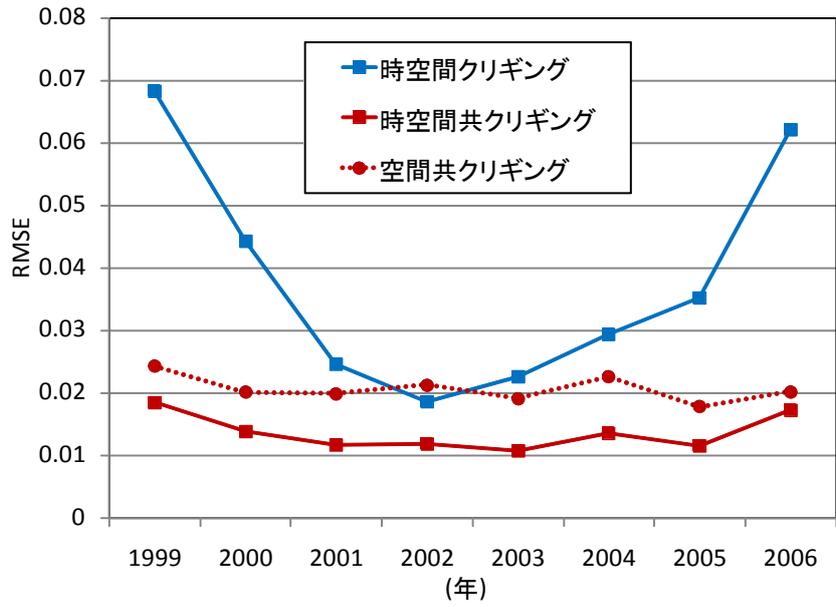


図3-8 年度毎の内挿精度 (RMSE)

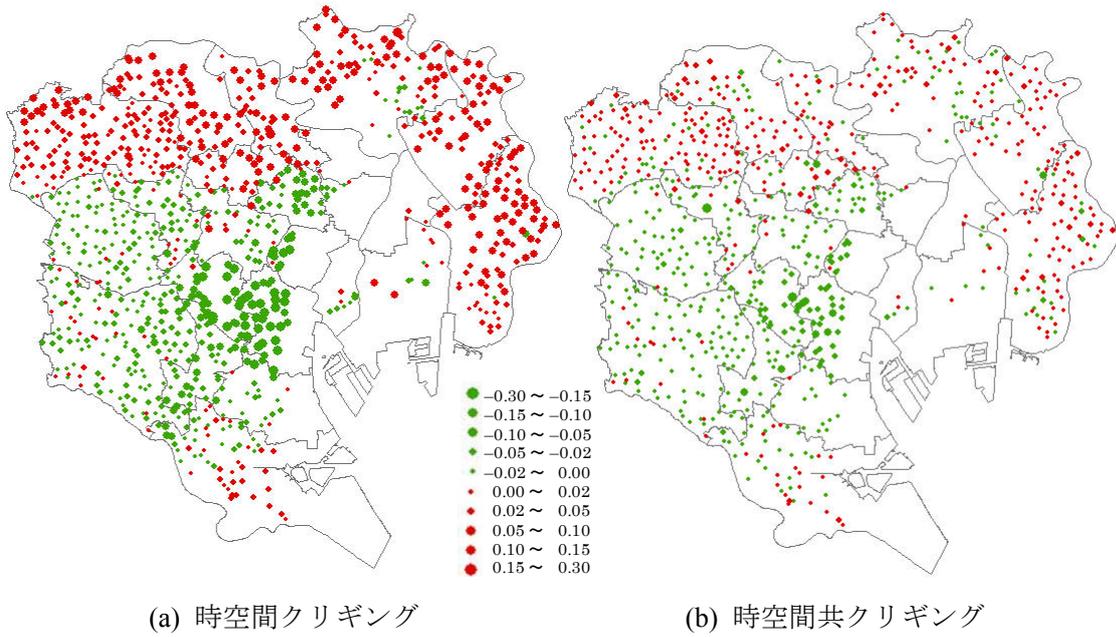
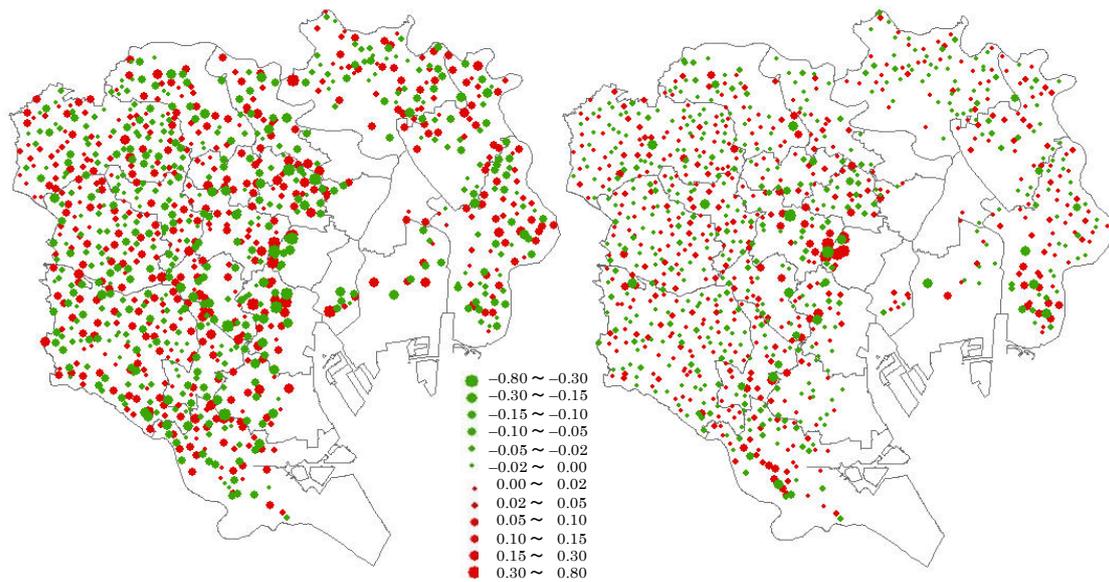


図3-9 内挿残差の空間分布 (2006年)



(a) 空間クリギング

(b) 空間共クリギング

図3-10 内挿残差の空間分布 (2006年)

4. 取引価格に関する情報提供

研究の背景で述べたように、不動産市場は、財の同質性や情報の完全性などが成り立たない典型的な不完全競争市場であるため、その動向把握には取引価格情報・鑑定価格情報の両面からの分析が必要不可欠である。全ての市場参加者が、これら 2 種類の価格情報の相互比較を通して市場分析を行える環境を整備することは、市場の透明性向上に大きく寄与するものとする。

しかし、既存の地価情報には、不動産市場の参加者に価格動向情報を提供する上で、必然的な限界が存在する。特に、新たに公開が開始された取引価格情報では、取引当事者の売り急ぎや買い急ぎなど個別事情を反映した情報が示されているため、必ずしも取引物件の標準的な価格を示さないなど問題が多く、現在公開されている取引価格情報だけから不動産市場の動向を把握することは難しい。

ここで、本研究でこれまで取り組んできた時空間で蓄積された鑑定価格情報の内挿というプロセスを踏むと、取引地点の鑑定価格情報を内挿し、取引価格と比較することが可能になる。すなわち、鑑定価格情報の時空間内挿を行うと、取引価格情報と鑑定価格情報の相互比較を通して市場分析を行える環境を提供することが可能である。

本章では、利用した地価データについて記した後、取引地点の推定と取引地点における鑑定価格情報の内挿について説明し、取引価格と内挿鑑定価格の比較による、地価情報の提供例と不動産市場分析例を示す。

4.1 利用した地価情報

取引価格情報は、国土交通省の土地総合情報システムで公開されている 2005 年第三四半期から 2008 年第二四半期まで 3 年間の住宅地・商業地・工業地の更地の取引価格、また、鑑定価格情報としては、1991～2008 年までの住宅系・商業・工業系用途の公示地価および基準地価を用いた。なお、データの制約から、路線価情報は使用していない。

また、取引価格データに含まれる取引件数は、10,333 件であるが、取引地点の推定や鑑定価格の内挿に支障のある取引を除外し、9,588 件の取引を用いた（表 4-1）。

表 4-1 除外した取引件数

データ件数	10,333
大字非公開	▲6
最寄り駅非公開	▲46
最寄り駅までの距離非公開	▲46
駅情報誤り	▲113
街区推定不可	▲281
容積率誤り	▲11
前面道路幅員無し、0m	▲242

4.2 取引地点の推定と鑑定価格の内挿

現在公開されている取引価格情報は、個人情報保護の観点から、個別取引の位置が特定されないように情報が隠されており、住所は大字単位までしか示されない。しかし、取引に対応した鑑定価格情報の内挿値を推定するためには、位置の情報が不可欠である。そこでまず、取引価格情報の属性から、「住所(大字)」「最寄り駅」「最寄り駅までの所要時間」を活用し、取引地点が存在する可能性のある範囲を推定した。

まず、国土院発行の「数値地図 2500(空間データ基盤)」から、街区区域・道路中心線・駅のデータを用いて各街区の最寄り駅および最寄り駅までの道路上距離を算出し、所要歩行時間を設定した。次に、取引価格情報の属性と、「住所(大字)」「最寄り駅」が一致し、かつ、「最寄り駅までの所要時間」が±1分以内の街区を抽出した。

図 4-1 は、取引存在範囲として推定された街区の総面積と取引地点として示された大字の面積の比率の分布を示している。全取引の約 70% が大字の 20% 以下の面積にまで範囲を絞り込めていることがわかる。一方、推定された街区の総面積が大字の面積の 90% 以上となっており、ほとんど絞り込みが出来ていない取引も存在する。このような取引には二種類あり、一つは大字自体が非常に小さく、全ての街区が前後 1 分の幅の中に入るもの（千代田区神田界限、中央区日本橋界限、新宿区神楽坂近辺）である。もう一つは駅からの所要時間が 30 分以上の取引である。公開の性質上、駅からの所要時間が 30 分を超えるものについては 30 分～60 分にまとめられているため、絞り込みによって除かれる街区がほとんど存在しない。このような取引は練馬区や足立区、葛飾区などに多く見られた。

本章では、以後この取引価格情報を用いて、鑑定価格情報の内挿を行い、取引価格との比較結果を示す。取引地点は取引存在範囲と推定された街区の重心を用いた。地価モデルは、取引価格情報の属性である「取引時点」「地積」「最寄り駅」「最寄り駅までの所要時間」「前面道路幅員」「容積率」を使用し、用途地域別にモデルを設定している。また、時空間共分散構造は球型バリオグラムで構造化している。

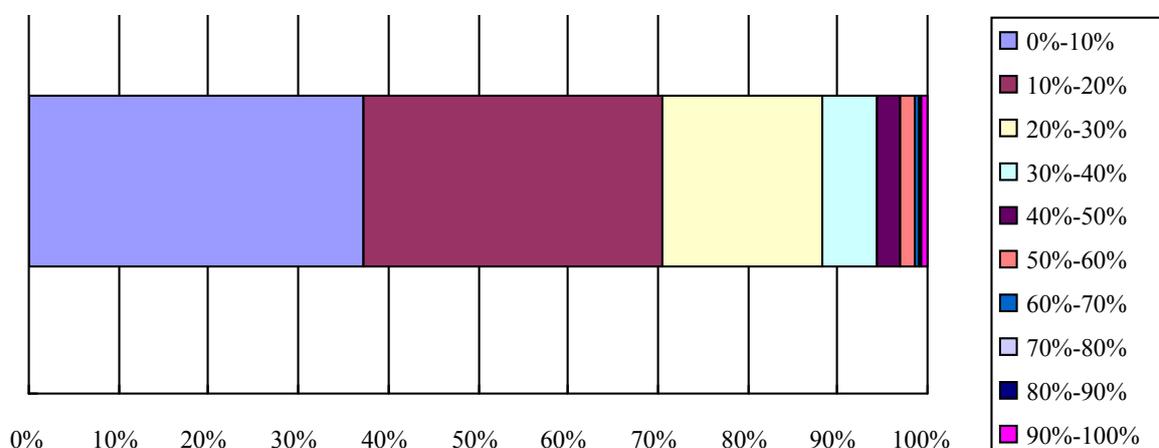


図 4-1 大字の面積に対する取引存在範囲の面積の割合

4.3 取引価格と鑑定価格の比較例

(1) 取引価格と鑑定価格の比較による地価情報の提供例

ここでは、現在の国土交通省が公開している地価情報提供サイト「土地総合情報システム」上に、取引価格情報を掲載することを想定した表示例を示す。

まず、図 4-2 はある 1 取引に関して、推定取引存在街区と対応する内挿価格を表示した例である。取引の存在範囲として推定された街区が地図上で赤く表示され、この取引に対応する鑑定価格の情報が別画面の棒グラフ上に表示されている。棒グラフ上に、緑点で推定取引存在街区の重心における鑑定価格情報を時空間内挿した値が示されており、赤点で取引価格が表示されている。この例は、鑑定価格情報と比較して極端に割安な価格で取引されたことを認識することができる。

次に、近隣の取引と比較し地域の市場動向を観察する目的で地価情報を閲覧する際の情報表示例を示す(図 4-3)。ここでは、取引存在範囲の重心に代表点を打ち、凡例のように鑑定価格の内挿値の最大値・平均値・最小値と取引価格の比較によって、点の色を変えて表示している。図 4-3 の点の位置は、必ずしも正確な取引位置を表現しているわけではないが、取引の空間分布や鑑定価格水準との比較結果が表現されており、近隣の不動産市場の動向を確認することができる。

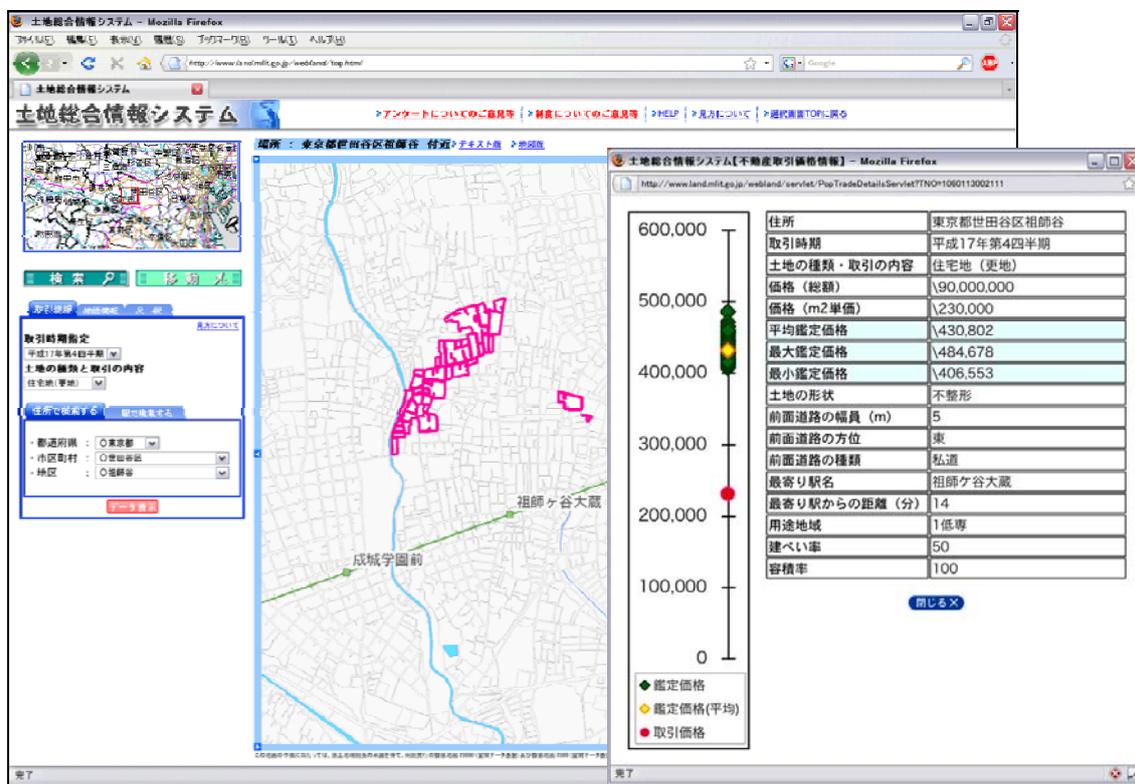


図 4-2 取引地点推定および鑑定価格内装を通じた取引価格情報の表示方法の提案

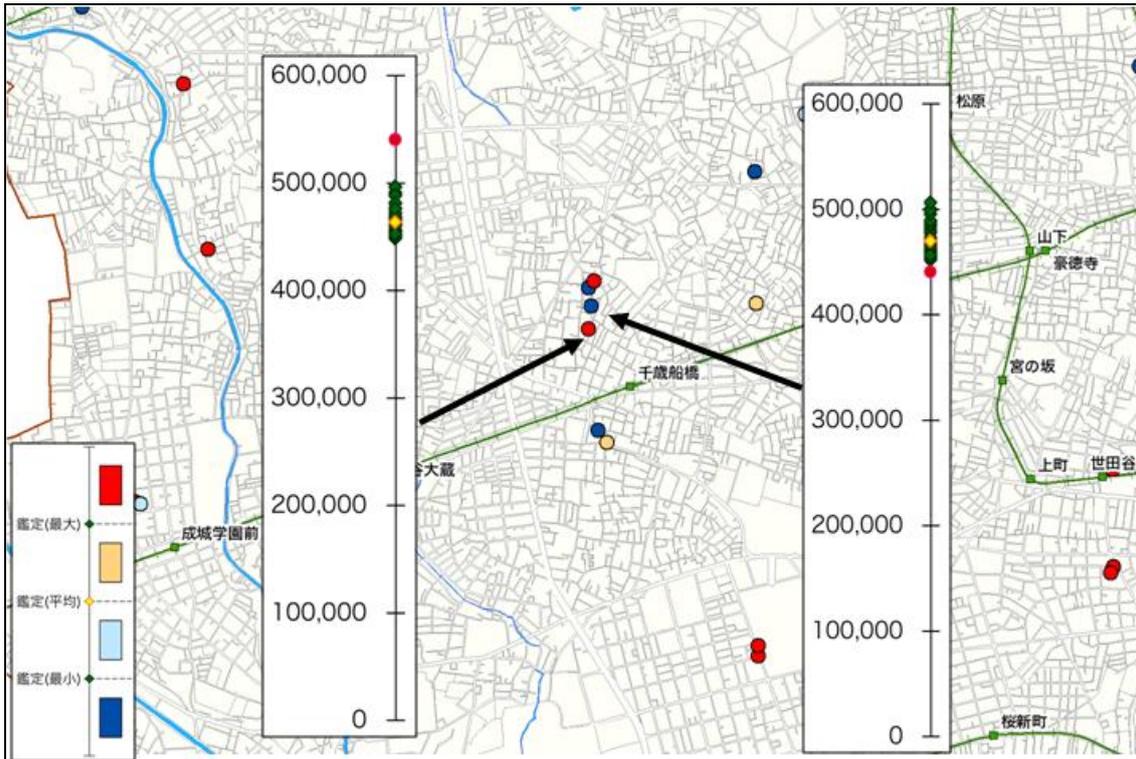


図 4-3 近隣取引の価格水準表現例（世田谷区祖師谷周辺）

(2) 取引価格と鑑定価格の比較による不動産市場分析

米国でのサブプライムローン問題に端を発した世界的な不況が日本にも暗い影を落していることは周知の通りである。ここでは内挿した鑑定価格情報と取引価格の比較を不動産市場の動向把握へ応用する例として、サブプライムローン問題の影響が日本に波及し始めた 2007 年第 4 四半期から 2008 年第 2 四半期にかけての住宅地の不動産市況を表した例を記す。

図 4-4～6 は 2007 年第 4 四半期から 2008 年第 2 四半期までの取引価格情報の様態を示したものである。上段の図は、推定された街区全体の重心に取引の位置を代表させた東京都 23 区全体の取引の分布図である。中段の図は鑑定価格に対する取引価格の大小で区ごとに件数を集計して円グラフに表したもので、円の大きさは取引の件数に比例させている。なお、取引価格情報が 10 件に満たない区については円グラフを表示していない。下段の図は横軸に鑑定価格の取引ごとの平均、縦軸に取引価格と鑑定価格の平均との比をとって取引をプロットしたものである。

2007 年第 4 四半期の時点では市場の変化は非常に小さく、全体的な分布はそれまでとほぼ同じ傾向である。南西にある世田谷区や目黒区では鑑定価格に比べて高い値段で取引された割合が高く、北東にある足立区や葛飾区では鑑定価格に比べて高い取引と安い取引の割合がほぼ等しくなっている。

2008 年の第 1 四半期になると 23 区全体の取引件数が約半数にまで減少する。細かく見る

と、取引の密度が特に薄くなっているのは葛飾区と目黒区・世田谷区にかかる東急線沿線であるが、前者は鑑定価格に比べ高い取引も安い取引も均等に減少しているのに対し、後者では鑑定価格に比べ高い取引のみが無くなっている。散布図を比較すると、鑑定価格の主要な価格帯である 500,000 円/m² 付近では分布に変化が見られず、均等に取引の密度が薄くなっていることがわかる。2007 年第 4 四半期で見られた、鑑定価格が 1,500,000 円/m² 付近で、かつ鑑定価格よりも高い取引は、2008 年第 1 四半期では姿を消している。これは高級住宅街である麻布や白金での取引が無くなったことに起因する。

2008 年第 2 四半期に入っても取引件数は依然として少ないままである。また、鑑定価格に比べ安く取引される割合が全体的に増加する。特に世田谷区・杉並区・目黒区などの西部では顕著に表れており、区ごとの円グラフで 2007 年第 4 四半期と比べるとその変化は明らかである。散布図で 2008 年第 1 四半期と比較すると、鑑定価格に対する取引価格の比が 0.8 を下回る取引が増加していることが分かる。また、このような取引は 23 区内で満遍なく散らばっている。公示地価・基準地価の分析では、今回取り上げた 3 期間では鑑定価格の水準はほとんど変化無しとしている。このことから、2008 年の第 2 四半期に入って取引価格は下落を始めたということが読み取れる。

4.4 まとめ

以上のように、地価の時空間相関構造を活かした鑑定価格情報の内挿と組み合わせることにより、取引価格情報と鑑定価格情報の比較を行うことが可能になる。

前述の通り、取引価格情報には取引従事者の個別事情の影響が多分に反映されている。しかし、本研究の提案である、鑑定価格情報の内挿価格との比較を通して取引価格情報を閲覧することにより、個別事情の影響をある程度把握することができるため、それらの事情に左右されずに不動産市場動向に関する情報を提供できると考えている。

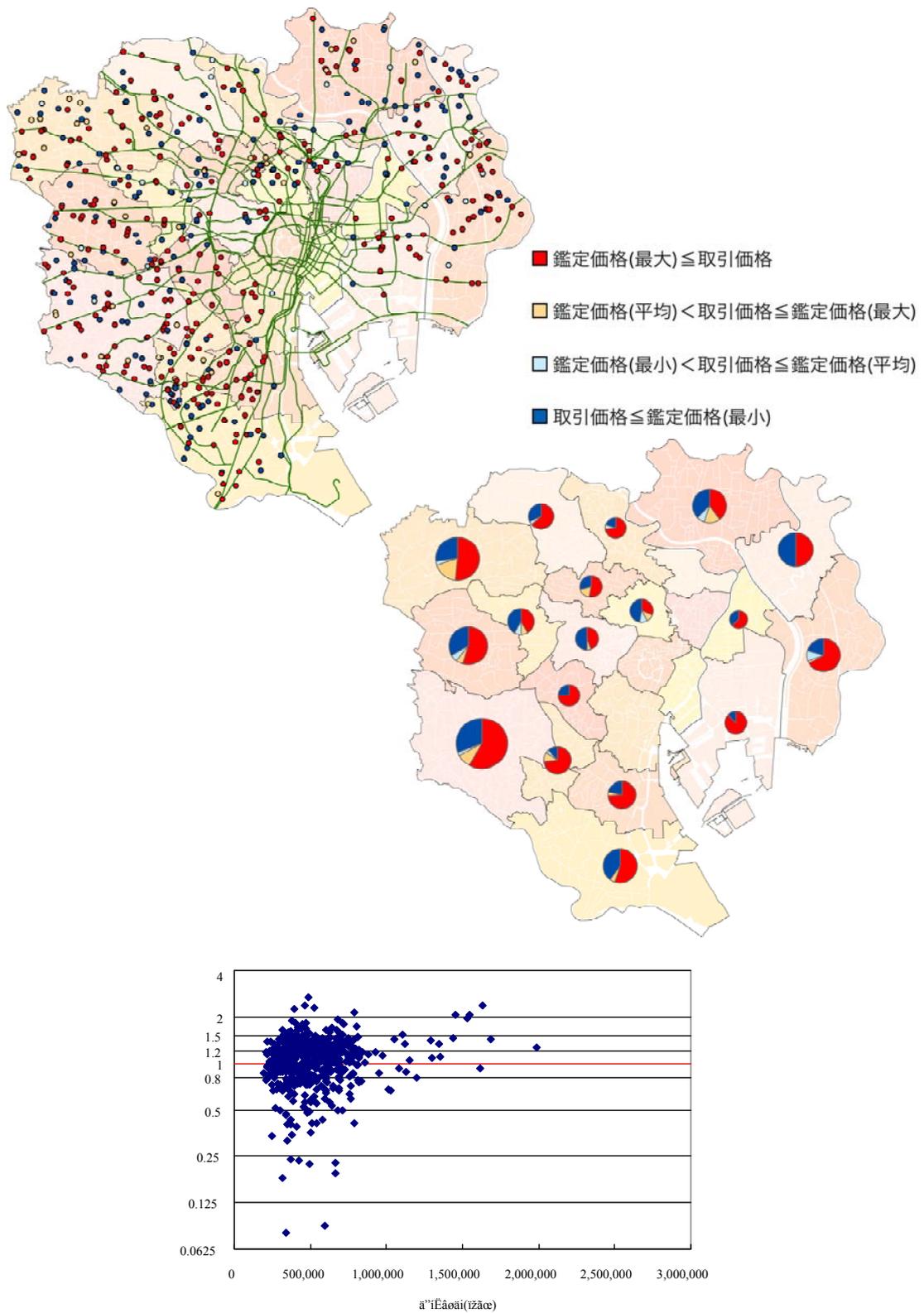


図 4-4 2007 年第 4 四半期の取引分布，区別割合，散布図

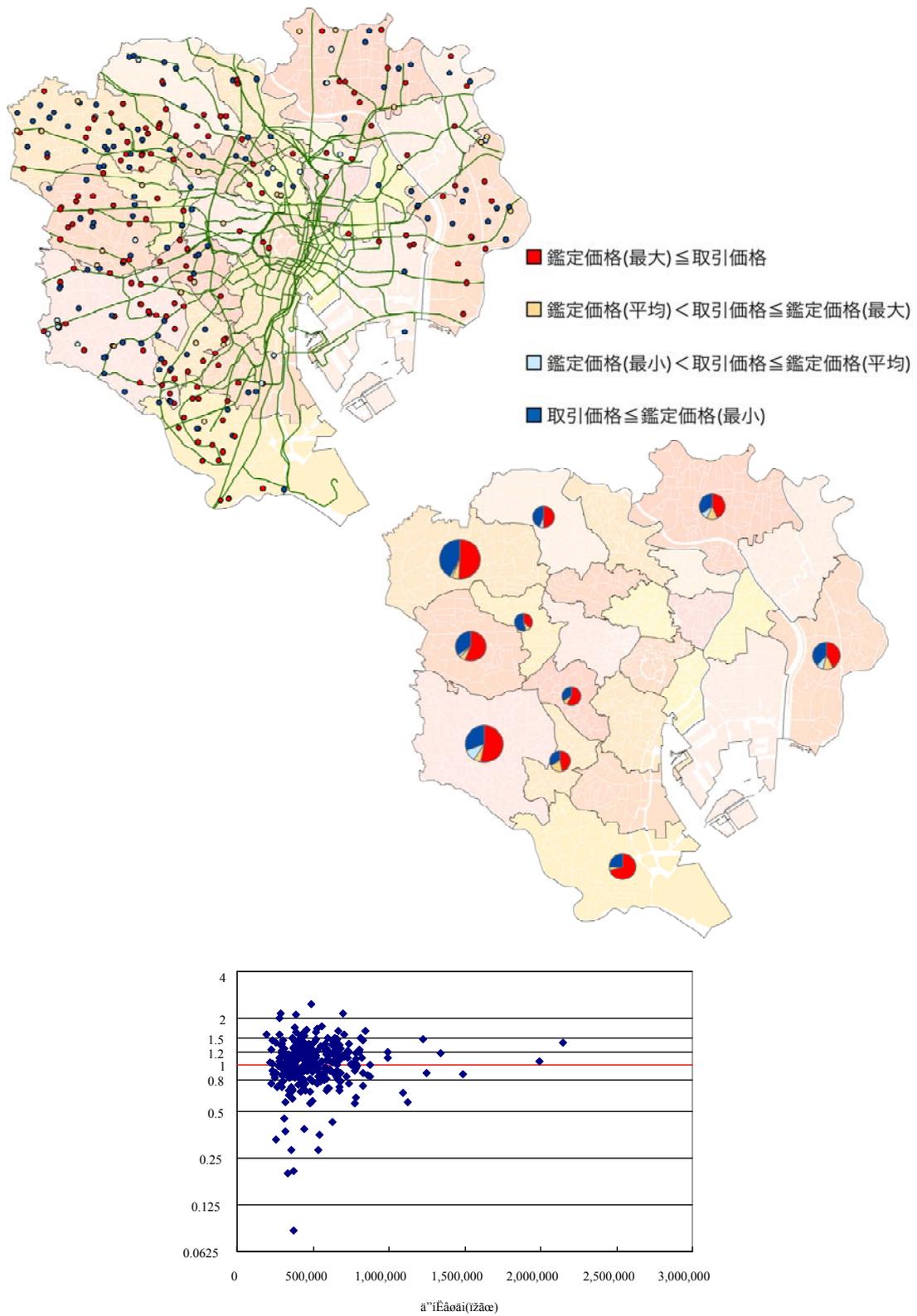


図 4-5 2008 年第 1 四半期の取引分布，区別割合，散布図

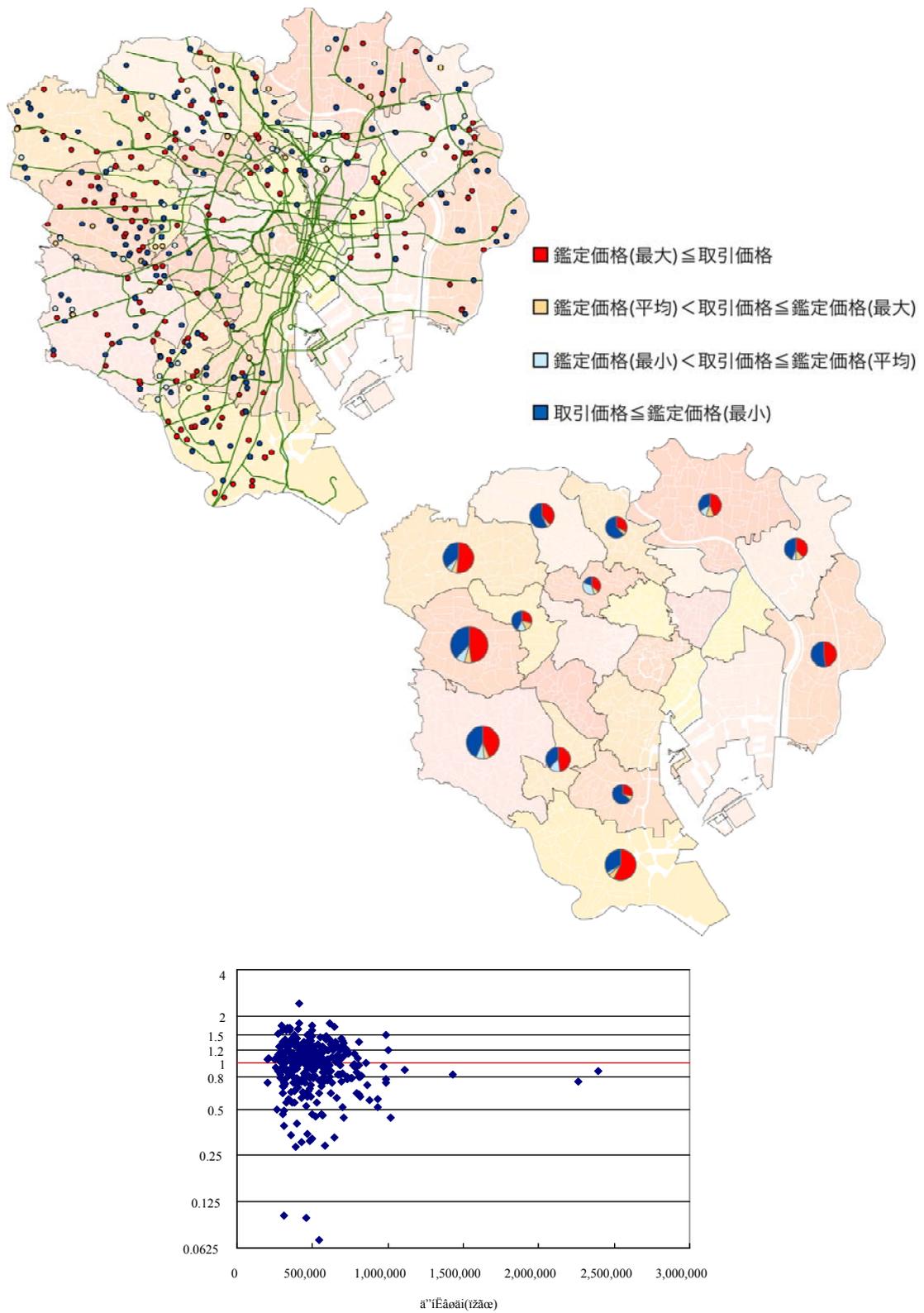


図 4-6 2008 年第 2 四半期の取引分布，区別割合，散布図

5. 結論

5.1 本研究の成果

不動産市場の透明性向上には、全ての不動産市場参加者が不動産市場の動向を把握できるだけの情報を提供する必要があるが、既存の地価情報には時空間での解像度が粗く、また鑑定価格・市場価格両面からの分析を行うには情報不足であるなど限界がある。そこで、本研究では、時空間で蓄積されてきた鑑定価格情報を活用した内挿という手法に着目し、任意の地点・時点での鑑定価格情報の提供と、市場取引価格との比較による市場分析が可能な環境の提供を目指した。

本研究では時空間内挿を行う手法として、空間統計学で議論されてきた共クリギングに着目し、公示地価情報の内挿に補助変数として路線価情報を活用して内挿精度の向上を図り、一定の適用可能性を確認した。更に、鑑定価格情報の時空間内挿に基づいて、取引価格情報・鑑定価格情報の比較による地価情報の提供を試みた。

本研究の主な成果を以下に整理する。

まず、共クリギングの地価内挿への適用可能性を、空間相関だけを考慮した実証実験で確認した。2006年の住居系用途地域内の公示地価を対象変数、相続税路線価を補助変数として、空間共クリギングによる公示地価の内挿を行い、内挿精度の評価を行った。追加補助変数（対象変数と異なる地点の補助変数）の空間分布パターンによる内挿精度への影響を、追加補助変数が一定間隔に分布するグリッド型・内挿地点のみに分布するコロケーション型による内挿精度の比較分析を通して評価した。実証実験の結果、共クリギングでは内挿地点から最寄りの補助変数までの距離が近いほど、内挿精度が高くなる傾向が見られ、グリッド型では追加補助変数の分布間隔が300mから100mに渡って距離短くなるほど内挿精度が上昇し、内挿地点に補助変数を与えるコロケーション型の内挿精度が最も高くなることが確認された。

次に、時空間相関を考慮した共クリギングに関する実証実験を行った。1999～2006年の住居系用途地域内の公示地価を対象変数、同期間の相続税路線価を補助変数とする時空間共クリギング（コロケーション型）を用いて公示地価の内挿を行い、内挿精度を時空間クリギングと比較した。その結果、全期間で内挿精度の向上が見られ、補助変数を利用する効果を確認できた。ただし、本実験では相続税路線価データの制約から短期間の情報しか利用することができず、時空間相関構造を正しく推定できていない可能性が示唆され、時空間相関を考慮した共クリギングの地価内挿への適用可能性については、今後更なる検討が必要である。

以上の実証実験の結果より、共クリギングの地価内挿への適用可能性について、空間相関については確認し、時空間相関については更なる検討が必要なものの内挿精度の向上を果たすことは確認できた。

また、地価の時空間内挿を利用した取引価格情報と鑑定価格情報の比較による情報提供

法を提案し、東京 23 区の取引価格情報を用いた例示を通して、不動産市場動向の把握に有効であることを示した。

以上のように、本研究では、地価内挿精度の向上に向けた共クリギングの適用可能性について一定の確認を行い、取引価格情報と鑑定価格情報の内挿を利用した不動産価格の動向に関する情報提供法の提案を行った。

5.2 今後の課題

今後の課題は以下の通りである。

まず、本研究では、地価内挿に対する時空間共クリギングの適用可能性に関して検討を行ったが、データ制約の問題から、長期間の情報を用いた分析を行うことができていない。今後、更なる検討が必要である。

また、共クリギングの実験では、補助変数として路線価情報を用いた。実験データの説明時に記載したように、路線価情報は公示地価と非常に強い相関を持っている。今後、より相関の弱い情報を補助変数として入力しても、共クリギングによる効果があるかどうか野検討が必要である。補助変数の候補としては、賃料などの不動産の収益性に関する情報などが考えられる。

更に、市町村・町丁目・メッシュなど一定の地域単位で集計された情報の補助変数としての利用可能性の検討も今後の課題である。例えば、住居系地価モデルでは所得水準や人口密度など住環境を表現する情報、商業系地価モデルでは従業者数や事業所数・商業売上高など商業地としての価値を表現する情報などは、地価と相関があると考えられ、補助変数としての利用の検討が望ましい。しかし、これまで空間統計学では、地点で取得された点データを用いた議論がなされており、面データを用いた検討はされていない。社会経済データへの応用を考えると、面データを用いた応用は極めて重要であると考えられる。

次に、情報提供に関する今後の課題であるが、第一は取引価格情報の生データを使った情報提供法の検討である。今回は、一般公開されている取引価格情報を用いた情報提供例を示したが、今後、取引位置が分かっている取引データを用いて、いかに有効な情報提供が可能か検討していきたいと考えている。

また、提案した鑑定価格情報と取引価格情報の比較を通じた不動産価格情報提供手法について、実装して公開することも今後の課題である。

参考文献

- Bishop, T. and Lark, R. (2007) "A landscape-scale experiment on the changes in available potassium over a winter wheat cropping season," *Geoderma*, Vol.141, pp.384-396.
- Burman, P. (1989) "A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods," *Biometrika*, Vol.76, No.3, pp.503-514.
- Chica-Olmo, J. (2007) "Prediction of housing location price by a multivariate spatial method: cokriging," *Journal of Real Estate Research*, Vol. 29, No. 1, pp.91-114.
- Chiles, J., and Delfiner, P. (1999) *Geostatistics Modeling Spatial Uncertainty*, Wiley-Interscience Publication: John Wiley & Sons, INC.
- Cressie, N. (1985) "Fitting variogram models by weighted least squares," *Mathematical Geology*, Vol.17, pp.563-586.
- Cressie, N. (1993) *Statistics for Spatial Data*, Wiley Series in Probability and Mathematical Statistics: John Wiley & Sons, revised edition.
- Cressie, N. and Huang, H. (1999) "Classes of nonseparable spatiotemporal stationary covariance functions," *Journal of the American Statistical Association*, Vol.94, pp.1330-1340.
- Geisser, S. (1975) "The predictive sample reuse method with application," *Journal of the American Statistical Association*, Vol.70, pp.320-328.
- Gneiting, T. (2002) "Nonseparable, stationary covariance functions for space-time data," *Journal of the American Statistical Association*, Vol.97, No.458, pp.590-600.
- Goldberger, A.S. (1962) "Best linear unbiased prediction in the generalized linear regression model," *Journal of the American Statistical Association*, Vol.57, No.298, pp.369-375.
- Goovaerts, P. (1997) "*Geostatistics for Natural Resources Evaluation*," Oxford University Press.
- Goulard, M. and Voltz M. (1992) "Linear coregionalization model: tools for estimation and choice of cross-variogram matrix," *Mathematical Geology*, Vol.24, pp.269-286.
- Journel, A.G. and Huijbregts, C.J. (1978) "*Mining Geostatistics*," Academic Press, London.
- Lark, R., Bellamy, P. and Rawlins, B. (2006) "Spatio-temporal variability of some metal concentrations in the soil of eastern England, and implications for soil monitoring," *Geoderma*, Vol.133, pp.363-379.
- Parado-Iguzquiza, E. and Atkinson, P. (2007) "Modelling the semivariograms and cross-semivariograms required in downscaling cokriging by numerical convolution-deconvolution," *Computers & Geosciences*, Vol.33, pp.1273-1284.

- Shao, J. (1993) "Linear model selection by cross-validation," *Journal of the American Statistical Association*, Vol.88, No.422, pp.486-494.
- Stone, M., (1974) "Cross-validation choice and assessment of statistical predictions," *Journal of the Royal Statistical Society*, Ser. B, Vol.36, pp.117-147.
- Wackernagel, H. (1998) *Multivariate Geostatistics*, Berlin Heidelberg: Springer-Verlag. (地球統計学研究委員会訳, 『地球統計学』, 森北出版, 2003 年)
- Webster, R. and Oliver, M. (2001) *Geostatistics for Environmental Scientists*, John Wiley & Sons, Ltd.
- Zhang, P. (1993) "Model selection via multifold cross validation," *Annals of Statistics*, Vol.21, No.1, pp.299-313.
- 井上 亮, 清水英範, 吉田 雄太郎, 李 勇鶴 (2009) 時空間クリギングによる東京 23 区・全用途地域を対象とした公示地価の分布と変遷の視覚化. GIS—理論と応用. 17(1): 13-24.
- 間瀬茂・武田純 (2001) , 『空間データモデリング—空間統計学の応用』, 共立出版.

SPATIO-TEMPORAL INTERPOLATION OF LAND PRICES USING COKRIGING

Inoue R.¹ Li.Y.¹ Shimizu.E.¹

¹Department of Civil Engineering, University of Tokyo

It is important to provide land price information based on real-estate appraisal and market transactions for all market participants to improve market transparency of real-estate market, however, present information does not have enough information. One way to provide land price information at any time and location is a spatio-temporal interpolation. In this study, we focus on cokriging; one of interpolation techniques discussed in Spatial Statistics, and empirically analyze its applicability for land price spatio-temporal interpolation. Experiments showed higher interpolation precision by cokriging than kriging. Then we propose a information prevision method on real-estate market through the comparison between market prices and appraised prices.

KEYWORDS: *land price, spatio-temporal interpolation, cokriging.*

研究 成 果 の 要 約

助成番号	助 成 研 究 名	研 究 者 ・ 所 属
第2008-9号	共クリギングによる土地取引価格の時空間内挿に関する研究	井上 亮・東京大学
<p>近年、市場原理によって土地の高度・有効利用を促進する施策の一環として、不動産市場の透明性の向上、特に地価情報の更なる整備と公開の必要性が叫ばれている。日本では地価情報の整備と公開の役割を、国土交通省による公示地価など不動産鑑定評価に基づく価格情報が担ってきたが、不動産市場での取引価格との乖離が常に指摘されてきた。そこで、国土交通省は土地の取引価格情報の公開方針を決定し、2006年4月から「土地総合情報システム」上で公開している。</p> <p>しかし、これら既存の地価情報は、時空間での解像度が粗く、また鑑定価格・市場価格両面からの分析を行うには情報不足であるなど、不動産市場の参加者に価格動向情報を提供する上で限界が存在する。そのため、市場参加者が関心を持つ土地の価格やその動向について十分な情報提供を行うことができない。そこで、本研究では、時空間で蓄積されてきた鑑定価格情報を活用した内挿という手法に着目し、任意の地点・時点での鑑定価格情報の提供と、市場取引価格との比較による市場分析が可能な環境の提供を目指した。</p> <p>本研究では、時空間内挿手法として、空間統計学で議論されてきた共クリギングを取り上げ、実証実験を通して適用可能性を検討した。まず、公示地価の内挿に補助変数として相続税路線価を活用して内挿精度の向上を図り、一定の適用可能性を確認した。更に、鑑定価格情報の時空間内挿に基づいて、取引価格情報・鑑定価格情報の比較による地価情報の提供を試みた。本研究の主な成果を以下に整理する。</p> <p>まず、地価内挿への共クリギングの適用可能性を、空間相関を考慮した実証実験で確認した。2006年住居系用途地域の公示地価を内挿対象変数、相続税路線価を補助変数とし、空間共クリギングの内挿精度を評価した。実験の結果、内挿地点から最寄りの補助変数までの距離が近いほど、内挿精度が高くなる傾向が見られ、補助変数を用いないクリギングより高精度の内挿が可能であることを確認した。また、等間隔に補助変数を配置するグリッド型では、補助変数に距離が近い内挿点の精度が上昇し、内挿地点に補助変数を与えるコロケーション型の内挿精度が最も高くなることが確認された。</p> <p>次に、時空間相関を考慮した共クリギングに関する実験を行った。1999～2006年の住居系用途地域の公示地価を対象変数、同期間の相続税路線価を補助変数とする時空間共クリギングを用いて公示地価の内挿を行い、内挿精度を時空間クリギングと比較した。その結果、全期間で内挿精度の向上が見られ、補助変数を利用する効果を確認できた。ただし、本実験では入力データの制約から短期間の情報しか利用することができず、時空間相関構造を正しく推定できていない可能性が示唆され、時空間相関を考慮した共クリギングの地価内挿への適用可能性については、今後更なる検討が必要である。</p> <p>以上の二実験より、共クリギングの地価内挿への適用可能性については、空間相関については確認し、時空間相関については更なる検討が必要なものの内挿精度の向上を果たすことは確認できた。</p> <p>また、地価の時空間内挿を利用した取引価格情報と鑑定価格情報の比較による情報提供法を提案し、東京23区の取引価格情報を用いた例示を通して、不動産市場動向の把握に有効であることを示した。</p> <p>このように、本研究では、地価内挿精度の向上に向けた共クリギングの適用可能性について一定の確認を行い、市場取引価格を用いた不動産価格の動向に関する情報提供法の提案を行った。</p>		