時空間Cokrigingの地価内挿への適用可能性の検討*

A Study on the Applicability of Spatio-Temporal Cokriging to Land Price Interpolation*

李　勇鶴**・井上　亮***・清水英範****

By Yonghe LI**・Ryo INOUE***・Eihan SHIMIZU****

１．はじめに

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、特に地価に関する情報のさら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もつ土地の価格や動向を知り、これを他の土地の情報と比較することができなければ、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、この「地価に関する情報の整備と公開」の役割をこれまで国土交通省による公示地価が担ってきた。地価公示点の数は以前に比べると増やされてきているが、全ての市場参加者が関心をもつ土地の価格や動向の情報を必ずしも提供できるわけではなく、現行の公示地価による情報提供には限界があるといわざるを得ない。従って、市場参加者が関心をもつ土地の価格や動向を知るためには、否応無しに時空間で蓄積された情報を利用した内挿というプロセスが不可欠になる。

さて、これまで地価の内挿には空間統計学の手法であるKrigingがよく利用されてきた。Krigingによる地価内挿では地価情報間の時間や空間上の相関を利用する[1]。しかし、もし内挿の対象となる地価情報（対象変数）と時間・空間上で強い相関のある他の情報（補助変数）が利用可能であれば、Krigingを拡張したCokrigingを用いることによって内挿精度をさらに高めることができる[2), 3), 4)]。しかし、現時点で、時空間Kriging[1)]や空間Cokriging[2)]による地価内挿の研究はみられるが、時空間Cokrigingを用いた研究はまだ行われていない。

そこで本研究では、対象変数と類似度が高く時間・空間上の相関が強い補助変数を利用して変数間の時空間相関を構造化し内挿を行う時空間Cokrigingに着目し、地価内挿への適用可能性を検討する。詳しくは、1999-2006年の東京23区の住宅地公示地価を対象変数とし、相続税路線価（以下は路線価とする）を補助変数として時空間Cokrigingの実証分析を行う。

* キーワーズ：時空間Cokriging、地価、時空間内挿

**非会員,工修,東京大学大学院工学系研究科社会基盤学専攻
(東京都文京区本郷7-3-1,TEL03-5841-6118,FAX03-5841-7453)
***正員,博(工),東京大学大学院工学系研究科社会基盤学専攻
(東京都文京区本郷7-3-1,TEL03-5841-6129,FAX03-5841-7453)
****正員,工博,東京大学大学院工学系研究科社会基盤学専攻
(東京都文京区本郷7-3-1,TEL03-5841-6126,FAX03-5841-7453)

２．時空間Cokriging

Cokrigingは、Krigingを拡張した手法で、多変量間の相関を利用して空間内挿を行う空間統計学の手法である。Cokrigingでは、特定の位置における対象変数の値を、近傍に存在しているその変数自体といくつかの補助変数に関するデータを基にして予測する[4)]。Cokrigingは、主に対象変数の数が少なく、自己相関だけを利用したKrigingでは高精度の予測ができない場合によく使われる。この時、対象変数と空間相関が強く観測数が多い補助変数を予測に利用するCokrigingを用いることによって、対象変数だけを利用したKrigingに比べて、予測精度を高めることができる。

Krigingでは自己共分散関数を用いて一変数の自己相関を構造化し空間内挿を行うのに対し、Cokrigingでは自己共分散に加えて異なる変数間の相互共分散も共分散関数を用いて構造化し、対象変数を空間内挿する。したがって、N個の変数を考慮したCokrigingではN×(N+1)/2個の自己・相互共分散関数を考慮しなければならなく、変数の数が大きくなるとかなり複雑になる。そのため、通常は二つの変数を用いたCokrigingをよく採用する。

N変量確率上において、2次定常性の仮定の枠組みで、自己・相互共分散関数$C_{ij}(\mathbf{h})$或いは自己・相互セミバリオグラム$\gamma_{ij}(\mathbf{h})$は以下のように定義される。

$$
\begin{aligned}
C_{ij}(\mathbf{h}) &= E[(Z_i(\mathbf{x}) - m_i)(Z_j(\mathbf{x}+\mathbf{h}) - m_j)] \\
\gamma_{ij}(\mathbf{h}) &= \frac{1}{2}E[(Z_i(\mathbf{x}+\mathbf{h}) - Z_i(\mathbf{x})(Z_j(\mathbf{x}+\mathbf{h}) - Z_j(\mathbf{x})]
\end{aligned}
\qquad (1)
$$

$$(\mathbf{x}, \mathbf{x}+\mathbf{h}) \in \mathbf{D} : i, j = 1 \cdots N)$$

Cokrigingでは、共分散関数を直接求めることではなく、まずバリオグラムを求めてから下記の共分散関数との関係式を用いて対応する共分散関数を求める。

$$C_{ij}(\mathbf{h}) = \gamma_{ij}(\infty) - \gamma_{ij}(\mathbf{h}) \qquad (2)$$

また、自己・相互共分散関数は半正定値符号関数であ

る必要がある。Krigingでは一つの共分散関数だけがこの条件を満たせばいいため、先行研究によって提案された関数（半正定値符号関数と検定された関数）を直接利用すればいいが、Cokrigingの場合はN×(N+1)/2個の関数から形成される自己・相互共分散関数組がこの条件を満たす必要があるため、これらの共分散関数はお互いに制約しており、独立に求めることはできない。それで、普段はLMCR（Linear Model of Coregionalization）を用いて半正定値符号条件を満たす自己・相互共分散関数を求める[3), 5)]。

地価データなどは空間上の相関だけではなく、時間上の相関も持っているため、時空間相関を考慮したCokrigingを行うとより高い精度の内挿が可能と考えられる。空間Cokrigingの時空間への拡張は自己・相互共分散関数で空間距離だけではなく時間差も考慮することで実現される。共分散関数を時空間に拡張する手法としては、時間軸を空間上の一つの軸と見なす手法、空間と時間上の共分散関数を独立に考慮する手法（Separable型）、空間と時間の相互作用を考慮する手法（Nonseparable型[6)]、Product-Sum型[7)]）などが挙げられる。

また、地価のような社会経済データでは、時空間距離の近さだけでは現象を説明することはできない。そこで、地価モデルの誤差に自己・相互共分散構造を仮定した時空間Universal Cokriging（以下は時空間Cokrigingとする）で地価の内挿を行うこととし、次章で住宅地公示地価と路線価を用いて地価内挿への時空間Cokrigingの適用可能性を検証する。

３. 実証実験

(1) 使用データと地価モデル

時空間Cokrigingの地価内挿への適用可能性を、1999-2006年の東京23 区の住宅地公示地価を対象変数、路線価を補助変数として実験を行った。対象地域・期間内には住宅地の公示地価データは8010点存在する。路線価データは、公示地価と同じ地点の路線価を使用することとし、最寄り道路の路線価格を設定した。但し、複数の道路と隣接している場合、最寄りの路線価を使う方法では間違った路線価が設定される場合がある。しかし、全地点の路線価をチェックすることは困難であるため、公示地価と比較して正常価格（公示地価の8割）より著しく低かった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。

また、本研究では公示地価と路線価の地価モデルの説明変数として、鉄道利便性指標(分)、最寄り駅までの道路距離(m)、容積率(%)、地積(㎡)、前面道路幅員(m)の情報を用いている。その中で鉄道利便性指標は、各地点の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）までの鉄道所要時間を各駅の乗降客数を元に重み付けして設定したものである。

地価モデルは、公示地価・路線価共に式(2)を用い、自己・相互セミバリオグラムは式(3)を用いた。なお、自己・相互セミバリオグラムのrange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[5)]を用いた。

$$\ln(P) = \beta_0 + \sum_{i=1}^{5} \beta_i \ln(x_i) + \varepsilon \qquad (3)$$

（但し、P: 地価(円/㎡)、x_1:鉄道利便性指標(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: 誤差）

$$\gamma(\mathbf{h}|\theta) = \tau^2 + {\sigma_1}^2 * Sph(-\|\mathbf{h}\|/\theta_1) + {\sigma_2}^2 * Sph(-\|t\|/\theta_2) \qquad (4)$$

（但し、γ：自己・相互セミバリオグラム、τ^2: nugget、σ_1^2, σ_2^2: 空間と時間上のpartial sill、θ_1, θ_2: 空間と時間上のrange、Sph: 球形型モデル）

時空間Cokrigingのパラメータ推定は、循環型のプロセスを採用した。まず、時空間 Krigingを用いて公示地価と路線価の残差を繰り返し計算[1)]で求め、初期値を設定する。次に、公示地価と路線価の残差を用いて経験自己・相互バリオグラムを計算し、それに理論自己・相互バリオグラムを当て嵌める[5)]。また、求まった理論自己・相互バリオグラムを利用して、時空間Cokriging方程式を用いて両地価モデルのパラメータを求める。パラメータが求まると、また公示地価と路線価の残差が求められ、新しい循環が始まる。この繰り返し計算はパラメータが収束するまで行う。

(2) 地価モデルと自己･相互セミバリオグラムのパラメータ推定結果

時空間Kriging(UK)と時空間Cokriging(UCK)の地価モデルのパラメータ推定結果を表―１に示す。UKとUCKの対象変数のパラメータを比較すると、前面道路幅員のパラメータだけがその絶対値がUCKでUKより小さくなり、他のパラメータはほぼ変わらない。ある説明変数のパラメータの絶対値がUCKで小さくなったのは、もともとこの説明変数によって説明されていた地価の一部がUCKで相互相関によって説明されたことを意味する。一方、UCK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比例関係を持っているためである。

セミバリオグラムのパラメータ推定には、両端だけの影響を避けるため　UKとUCKの自己・相互セミバリオグラムのパラメータ推定結果を表―２に示した。表―２から分かるように、四つのセミバリオグラムは類似しており、時空間相関構造は似ていると考えられる。それで、本研究ではその中からUCK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、詳細分析を行った（図―

１）。図―１の経験セミバリオグラムをみると、nuggetは0.01ぐらいで非常に小さい。それに比べて、空間軸の経験セミバリオグラムは15kmで0.14ぐらいになりnuggetとの比が14もあるため、空間上の相関はかなり強いことが分かる。一方、時間軸の経験セミバリオグラムはnuggetと比べてほぼ変化がないため、時間上の相関は弱いとみられる。経験セミバリオグラムを元に求められた理論セミバリオグラムは図―１のようである。

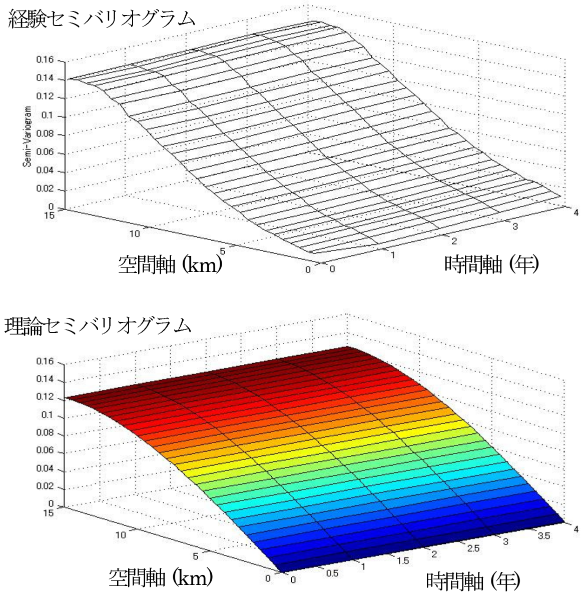
表―１　地価モデルのパラメータ推定結果

	UK (対象)	UCK (対象)	UCK (補助)
β_0	13.886 (33.7)	13.945(35.1)	13.749 (34.8)
β_1	-0.168 (-5.0)	-0.177 (-6.1)	-0.171 (-6.0)
β_2	-0.075(-13.2)	-0.074 (-15.3)	-0.074 (-15.4)
β_3	-0.081 (-8.3)	-0.082 (-9.8)	-0.083 (-10.0)
β_4	0.081 (5.1)	0.065 (4.2)	0.064 (4.2)
β_5	0.047 (4.8)	0.047 (5.1)	0.040 (4.4)

表―2　自己・相互セミバリオグラムの
パラメータ推定結果

	UK(γ_{11})	UCK(γ_{11})	UCK (γ_{22})	UCK (γ_{12})
τ^2	0.0001	0.0001	0.0001	0.0001
σ_1^2	0.115	0.123	0.122	0.122
θ_1	15	15	15	15
σ_2^2	0.001	0.001	0.001	0.001
θ_2	4	4	4	4

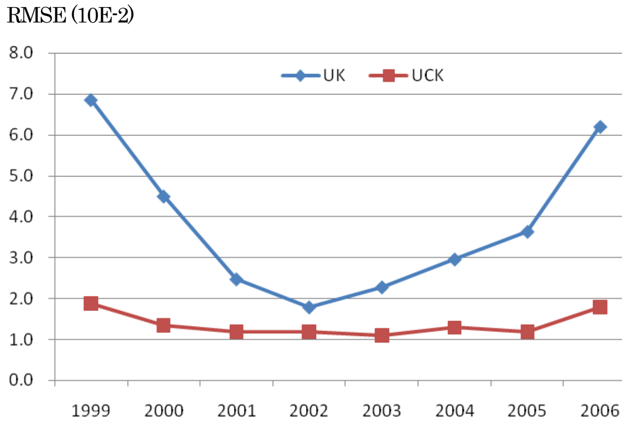
（注 : $\gamma 11$, $\gamma 22$: 対象変数と補助変数の自己セミバリオグラム、
$\gamma 12$:　相互セミバリオグラム）



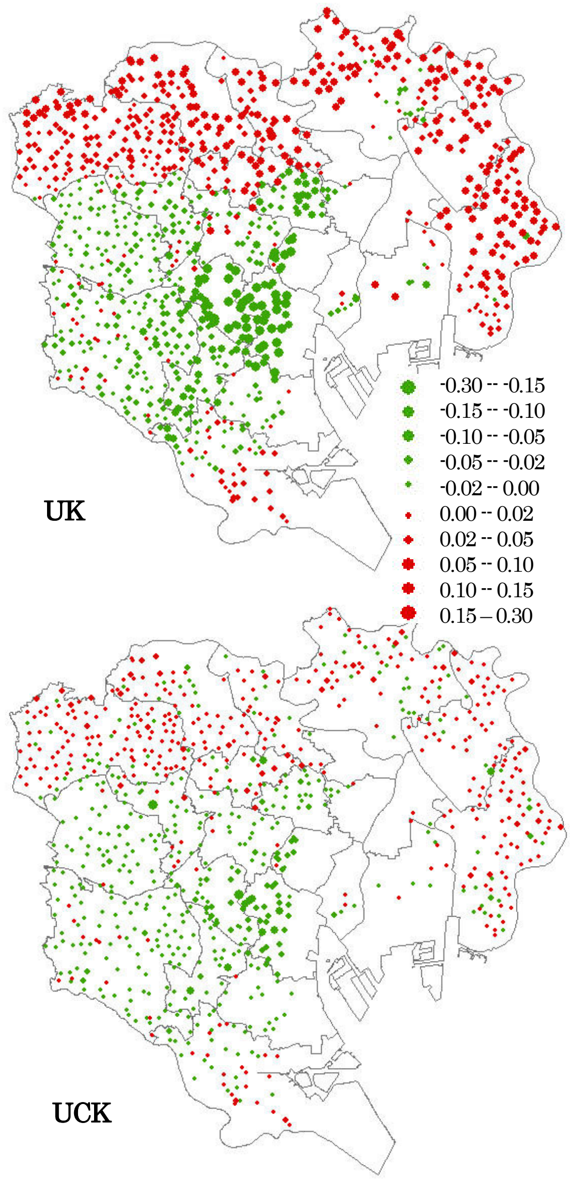
図－１　対象変数の経験・理論自己セミバリオグラム

(3)　UKとUCKの内挿精度の比較

UKとUCKの内挿精度を5-fold cross validationを用いて比較した。5-fold cross validationでは、データを無作為に五つのクループに分け、そのうちの一つのグループを精度検証用とし、残る四つのクループをモデル構築用とする。また、五つのクループそれぞれを一回ずつ精度検証用として5回繰り返す。但し、クループ化したのは対象変数だけで、補助変数はすべての検証で全部のデータを用いた。つまり、内挿地点の補助変数は分かっていることになる。



図－２　年毎の内挿精度



図－３　内挿残差の空間分布（2006年）

検証の結果、内挿精度の評価指標であるRMSE (Root Mean Square Error)は時空間Krigingの0.044に対し、時空間Cokrigingでは0.018となっており、時空間Cokrigingによる内挿精度が優れていることが分かる。

年別の内挿精度を図―２に示したが、全期間で時空間Cokrigingの内挿精度は時空間Krigingより高い。また、時空間Krigingの内挿精度は2002と2003年の対象期間の中間で最も高く、両端（1999年と2006年）では大きく低下する。一方、時空間Cokrigingでは全年度で同じレベル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Cokriging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に、2006年の内挿残差（ln(予測価格)-ln(鑑定価格)）を地図上にプロットした（図―３）。まず、内挿残差の大きさをみると、UKでは中心部と郊外部の東側と北側で内挿精度がかなり悪いが、UCKでは中心部の数箇所を除いて全域で高精度の内挿ができていることが確認できる。さらに、UCKの内挿精度がかなり悪い地点（内挿残差の絶対値が0.1より大きい地点）をみると、そのすべてが鑑定価格より低く内挿されている。その原因は、これらの地点は複数の道路に隣接しており、路線価を設定した最寄りの道路が実は全面道路ではなくて、正確な路線価より低い間違った路線価が付いていたためである。したがって、正確な路線価に修正すれば内挿精度もさらに向上すると考えられる。次に、内挿残差の正負をみると、UKでは郊外部の東側と北側（赤いポイント）で鑑定価格より高く予測され、残りの地域（青いポイント）では低く予測されたことが分かる。その原因は、Krigingでは定常性を仮定してすべての地点で地価の変動が同じであるとみなすが、実際この両地域での地価の変動は異なるためである。同じ現象は時空間Cokrigingでも現れているが、時空間Krigingよりは若干弱くなっていることがわかる。

以上のように、時空間Cokrigingを地価内挿に活用することにより、内挿精度の向上を図ることができることが実験より確認された。

４．おわりに

本研究では、時空間予測を類似度の高い補助変数を利用して変数間の時空間相関を構造化し内挿を行う時空間Cokrigingに着目し、1999-2006年の東京23区における住宅地公示地価データと路線価データを用いて実証実験を行い、時空間Cokrigingの地価内挿への適用可能性を時空間Krigingと比較して検証した。

実証実験の結果、時空間Cokrigingの内挿精度（RMSE）は時空間Krigingの4割程度で著しく向上したことが明らかになった。年次毎の内挿精度を比較してみたところ、時空間Krigingでは内挿精度が対象年度の両端（1999年と2006年）で急激に下落したが、時空間Cokrigingでは全年度での高精度の内挿ができたことが分かった。また、内挿精度の空間分布は、時空間Krigingでは中心部と郊外部の東側と北側でかなり悪かったが、時空間Cokrigingでは中心部の数箇所を除いて全域で高精度を保っていた。他に、時空間Krigingと時空間Cokrigingは両方とも郊外部の東側と北側で高く予測されており、残りの地域で低く予測されていたが、時空間Cokrigingではその傾向が若干弱かった。

今後の課題としては以下が挙げられる。

図―３から分かるように、時空間Cokrigingの内挿残差はまだ空間上の相関を持っている。但し、その相関は二つの地域で異なるため、現在の手法ではその相関を構造化することができない。それで、この現象に対する時空間相関の構造化手法を提案することによって、時空間Cokrigingの内挿精度をさらに向上させることが期待される。

また、本研究では路線価データの制限で1999-2006年の8年間のデータだけを用いて実験を行ったため、経験バリオグラム（図―１）から時間上の相関を正確に読み取ることができなかった。そのため、今後、より長い期間のデータを用いて実験を行う必要があると考えている。

参考文献

1) 井上亮，木越尚之，清水英範：時空間クリギングの地価推定への適用可能性の検討，地理情報システム学会講演論文集，No.14，pp.39-42，2005.

2) Jorge Chica-Olmo : Prediction of housing location price by a multivariate spatial method: cokriging，Journal of Real Estate Research，Vol. 29，No. 1，pp. 91-114，2007.

3) 本多 真, 菊地宏吉, 鈴木哲也, 水戸義忠：ダム周辺地下水位変動の時空間解析のための時空間Cokrigingの開発と適用，土木計画学論文集，No.659，pp.283-295，2000.

4) Hans Wackernagel: Multivariate geostatistics: An introduction with applications，Springer-Verlag，Berlin，2003［青木謙治監訳，地球統計学研究委員会訳，地球統計学，森北出版株式会社，2003］

5) Goulard, M. and Voltz, M.: Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix, Mathematical Geology, Vol. 24, No. 3, pp.269-286, 1992.

6) Cressie, N. and Huang, H-C: Classes of Nonseparable Spatio-Temporal Stationary Covariance Functions, Journalof the American Statistical Association, Vol.94, No.448, pp.1330-1340, 1999.

7) De Iaco, S., Myers, D.E., Posa, D. : Space-Time analysis using a general product-sum model, Statistics & Probability Letters, Vol.52, pp.21-28,2001.