応用力学論文集 Vol. 3, pp. 125-132 (2000 年 8 月)　　土木学会

# 誤差要素モデルに基づく Kriging を用いた空間内挿

# Spatial Interpolation by Kriging based on Error Component Model

堤盛人*・清水英範**・井出裕史***
Morito TSUTSUMI, Eihan SHIMIZU and Hiroshi IDE

*正会員　博士（工学）　東京大学講師　大学院工学系研究科　社会基盤工学専攻
**正会員　工博　東京大学教授　大学院工学系研究科　社会基盤工学専攻
（〒113-8656　東京都文京区本郷 7－3－1）
***正会員　修士（工学）　日本政策投資銀行　本店都市開発部
（〒100-0004　東京都千代田区大手町 1－9－1）

Among some sophisticated interpolation techniques, Kriging is an attractive method because it is constrained by intrinsic hypothesis. In order to apply Kriging to actual data, however, it is required to choose the variogram or covariance function, which are representation of spatial correlation. It is well known that spatial autocorrelation of error terms in regression model is often encountered, for which error component models (ECM) have been suggested and applied. In this paper, an interpolation technique coherent with Kriging system by expanding ECM is developed.

***Key Words : spatial interpolation, spatial statistics, spatial econometrics, Kriging, error component model***

## 1. はじめに

地理的・空間的に連続して分布するデータ（空間データ）を扱う分析においては、結果の表示や分析自体のために、何らかの内挿を必要とする場合が少なくない。空間データを合理的に内挿する手法の一つとして、鉱物学や地質学など自然科学と関わりの深い空間統計学の分野においては Kriging と呼ばれる手法が広く知られている。Kriging は、ある条件下で最良線形不偏推定量を与えるという意味で統計学的に優れた手法である。Kriging の適用に際しては、データの共分散に関して共分散関数若しくは variogram と呼ばれるものの情報が必要となる。

一方、社会経済分析と関わりの深い空間計量経済学と呼ばれる分野においては、空間データを用いた回帰分析において誤差項に何らかの系列相関が生じる場合に、誤差要素モデル(Error Component Model : ECM)と呼ばれるモデルが用いられる。ECM は、誤差項における未知の空間的相関に対し、パラメータを用いたモデル化によって対処しようとするものである。

本論文では、ECM に基づきながら Kriging を適用して空間内挿を行うことの理論的課題とその解決法について検討を行う。まず第 2 章において、ここで対象とする空間内挿の手法の一つであり、空間統計学において広く知られて用いられている Kriging について、その基本的考え方と理論の前提を概説する。続く第３章では、内挿とは別に、空間の相互作用を考慮した回帰分析の方法の一つとして、空間計量経済学等の分野で用いられている ECM について簡単に説明したうえで、ECM に基づき、Kriging と整合する空間内挿の手法について理論的に考察する。そして、この問題点を克服するための方法論について第４章で提案を行い、方法の妥当性を吟味する。

## 2. Kriging による空間内挿

1.で述べたように、地理的・空間的に連続して分布するデータ（空間データ）を扱う分析においては、分析を行うために、あるいは得られた結果を分かり易く表示するために、何らかの方法による内挿を必要とする場合が少なくない。内挿空間データを合理的に内挿する手法に関して、多項式の当てはめによる spline 補間などに代表される多くの手法が提案され、実際に様々な場面で適用されている。そのような内挿手法の一つとして、鉱物学や地質学など自然科学と関わりの深い分野である geostatistics （地球統計学などと訳される場合もある）においては Kriging が広く知られており、実際の分析・調査に利用されている（例えば、Cressie(1993) [1], Bailey and Gatrell (1995) [2]）。さらに、近年では、これらの空間を対象とした統計分析の分野を包含する形で空間統計学

spatial statistics と呼ばれる分野が形成されつつあり[3]、その進展が注目されている。

Kriging に関しては、今なお様々な拡張が試みられているが、本章では以降の議論に必要な部分に限って手法の概説を行う。

### 2.1 Kriging の基本的考え方

Kriging では、データが未知の地点 $k$ における値 $\hat{z}_k$ が、周囲のデータが既知の観測点における値 $z_i$ の線形和として、次式のように書けるものと仮定する。

$$\hat{z}_k = \sum_{i=1}^{I} \lambda_i z_i = \boldsymbol{\lambda}^t \boldsymbol{z} \tag{1}$$

$t$ はベクトル及び行列の転置を表し、 $i$ は観測点の番号を表す（ $i=1,2,\cdots,I$ ）。

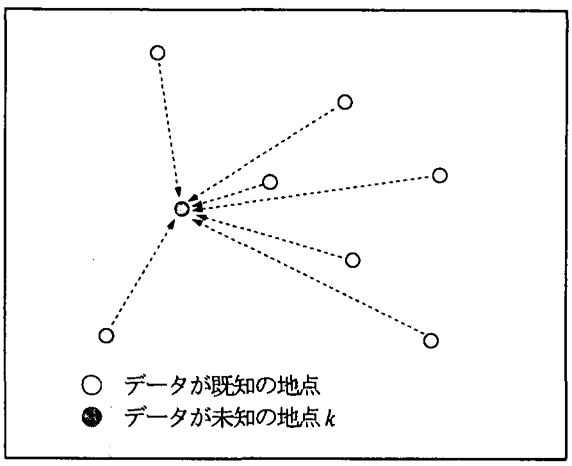


図１．Kriging による内挿の概略

観測点におけるデータ相互の共分散行列 $\boldsymbol{C}$ と、内挿により値を推定したい地点のデータと観測点におけるデータとの間の共分散ベクトル $\boldsymbol{c}$ が与えられたもとで、最良線形不偏推定量（Best Linear Unbiased Predictor : BLUP）を与えるのが Kriging である。具体的には、不偏性に関する制約条件のもとでの期待二乗誤差最小化から、式(1)における重み$\lambda_i$が次のように導かれる。

$$\boldsymbol{\lambda} = \boldsymbol{C}^{-1}\boldsymbol{c} \tag{2}$$

ここで、 $\boldsymbol{\lambda}$ は $\lambda_i$ を各要素とするベクトルである。

共分散行列 $\boldsymbol{C}$ や共分散ベクトル $\boldsymbol{c}$ の与え方にはいくつかの方法があるが、最もよく用いられる方法の一つは、$z_i$ , $z_j$ の共分散が次式のように距離 $d_{ij}$ のみによって決まると仮定するものである。

$$Cov(z_i, z_j) = c(d_{ij}) \tag{3}$$

$c(d_{ij})$ は共分散関数（covariance function）と呼ばれる。ここで、式(3)に示すように、任意の２地点 $i,j$ における値 $z_i$, $z_j$ に関する共分散がそれら 2 地点間の距離 $d_{ij}$ のみによって定義されるために、以下に示すような定常が仮定されている。

$$E(z_i - z_j) = 0 \tag{4}$$

$$Var(z_i - z_j) = 2 \cdot \gamma(d_{ij}) \tag{5}$$

式(5)における関数 $\gamma(\cdot)$ は variogram（詳しくは semi-variogram）と呼ばれ。そして、これに対応して、共分散関数 $c(d_{ij})$ は covariogram と呼ばれることがある。

定常の概念は本論文において最も重要な部分である。内挿に関わる重みのパラメータ $\boldsymbol{\lambda}$ は、不偏性の制約条件付きの期待二乗誤差最小化から導出されるが、その際、この定常の仮定を用いることで式(2)のような解が導出される。逆に、定常性が満足されない場合に式(2)を用いて内挿を行うと、推定結果にバイアスが生じることとなる。一方、次章で説明する誤差要素モデルでは、この定常性（特に２次モーメントに関する仮定：式(5)）を仮定しないため、結果として一般にはこれが満足されないこととなる。

なお、式(4)及び(5)では確率変数の差のみに対する定常性の概念であり、厳密には弱定常（weak stationary）と呼ばれる。これは、2.2 で示すような、確率変数そのものに対してより強い仮定をおく場合の定常（強定常）とは一般には区別されるが、以後本論文では両者を特に区別せず定常と呼ぶこととする。

式(5)を式(4)を用いながら展開すると、次式を得る。

$$2 \cdot \gamma(d_{ij}) = Var(z_i) + Var(z_j) - 2Cov(z_i, z_j) \tag{6}$$

一方、$z_i, z_j$ の（自己）分散は次式のように表される。

$$Var(z_i) = Var(z_j) = c(0) \tag{7}$$

式(7)と式(3)を式(6)に代入することにより次式を得る。

$$\gamma(d_{ij}) = c(0) - c(d_{ij}) \tag{8}$$

式(2)を用いて Kriging を実際のデータに適用するためには、$\boldsymbol{C}$ と $\boldsymbol{c}$ が既知である必要があり、通常は共分散関数（*or* covariogram）$c(\cdot)$、若しくは式(8)に示されているようにこれと直接関係する variogram $\gamma(\cdot)$のいずれか一方を決定する必要がある。

### 2.2. Universal Kriging

以下、本論文では、$y$ を被説明変数、$x_h$（ $h=1,2,\cdots,H$）を説明変数、$\beta_h$ （ $h=1,2,\cdots,H$）をパラメータとする線形回帰モデルを扱う。

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_H x_H \tag{9}$$

そこで、次に、この回帰モデルの推定結果に Kriging を適用する方法について簡単に説明する。誤差項を含んだ統計モデルを以下のように行列表記する。

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{10}$$

$$y = (y_1, \cdots, y_i, \cdots, y_I)', \quad \beta = (\beta_0, \beta_1, \cdots, \beta_H)', \quad X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{H1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1I} & \cdots & x_{HI} \end{pmatrix}, \quad \varepsilon = (\varepsilon_1, \cdots, \varepsilon_i, \cdots, \varepsilon_I)'$$

$\boldsymbol{\varepsilon}$は誤差項であり、通常は、式(11),(12)を満たすものと仮定する。

$$E(\varepsilon) = \boldsymbol{0} \tag{11}$$

$$Var(\varepsilon) = \sigma^2 \boldsymbol{I} \tag{12}$$

誤差項に対する通常の仮定（式(12)）からの違背が生じる場合でも、誤差項の共分散行列 $Var(\varepsilon)$が既知であれば、一般化最小二乗法（Generalized Least Squares Method : GLS）などによりパラメータを推定すればよい。例えば、$Var(\varepsilon)$ が既知の行列 $\boldsymbol{\Sigma}$ で表されるとすると、一般化最小二乗法によるパラメータの推定値（量）は次式により求められる。

$$\boldsymbol{\beta}_{GLS} = \left(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \tag{13}$$

誤差項 $\boldsymbol{\varepsilon}$ は確率変数であり、これを2.1における $z$ と見ることにより、観測地点以外の地点の $y$ の値をKriging によって推定することができる。具体的には、パラメータ $\boldsymbol{\beta}, \boldsymbol{\lambda}$ が求められれば、ある地点 $k$ における推定値（量）$\hat{y}_k$ は次式により求められる。

$$\hat{y}_k = \boldsymbol{x}_k \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\lambda}}' \hat{\boldsymbol{\varepsilon}} \tag{14}$$

（$\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\varepsilon}}$ ：$\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\varepsilon}$ の推定値（量））

ただし、$\boldsymbol{x}_k$ は$\boldsymbol{X}$の $k$ 行目のベクトル$(1, x_{1k}, x_{2k}, \cdots, x_{Hk})$である。

ここで、確率変数 $\boldsymbol{\varepsilon}$ に対し、式(8)における共分散行列 $\boldsymbol{\Sigma}$ と式(2)における共分散行列 $\boldsymbol{C}$ が全く異なるという状況を想定することも、技術的には何ら問題無く可能であるが、$\boldsymbol{\Sigma}$ と $\boldsymbol{C}$ は同じ共分散行列であると考える方がより合理的である。

上述のような考え方に基づく内挿のモデル化は、Universal Kriging と呼ばれる。本論文では、以後、このUniversal Kriging の適用を前提とする。

## 3. 誤差要素モデルと空間内挿

クロス・セクションデータに対して回帰分析などを行う場合、データの測定問題やモデルから抜け落ちたデータ間の外部効果や波及効果のためにデータサンプルの誤差項同士に相関が残る、すなわち空間的自己相関（Spatial Autocorrelation）が存在することがある。また、空間相互作用や空間拡散過程といった空間過程（Spatial Process）をモデルが十分に説明しきれない場合、近隣ゾーンの被説明変数間にある関数関係が存在することがある。Anselin (1988)ら [4]は、それらを総称して空間的依存性（Spatial Dependence）と呼んでいる。

周知のように、誤差項に系列相関が存在する場合、これを無視して通常最小二乗法(Ordinary Least Squares Method: OLS)による推定を行うと、推定量が有効性を持たなくなるといった問題が生じる。また、OLS を用いた場合の分散推定値が偏りをもつため、通常の有意性検定は無効になってしまう。

### 3.1 誤差要素モデル

空間を対象としたモデルの定式化にあたっても、明らかに空間依存性を引き起こすと考えられる要因はモデルに取り込まれていることが望ましいのは言うまでもない。しかし、実際にはそれらの要因をすべて考慮することは不可能であり、結果として残差に空間的自己相関が生じることが多い。このような誤差項に対する通常の仮定からの違背が生じる場合でも、その構造が既知であれば、2.1 で述べたように一般化最小二乗法などによりパラメータを推定すればよい。しかし、実際にはその構造が既知であることは稀であり、一般にはそれを何らかの仮定に基づきモデル化する必要が生じる。

ところで、系列 series という言葉はもともと（通例３つ以上の）一続き・一組を意味するが、系列相関 serial correlation のうち統計学において最も盛んに研究が行われてきたのは言うまでもなく時系列 time series である。その処理に関しては、自己回帰（autoregressive : AR）モデルや移動平均（moving average : MA）モデルなどが良く知られている。空間計量経済学や計量地理学の分野においては、これら時系列分析におけるアプローチのアナロジーから、誤差項の空間的自己相関に対する対処法として次のようなモデルが提案されている [4),5)]。

(a) 自己回帰モデル : Auto-Regressive Model

$$\boldsymbol{\varepsilon} = \rho \boldsymbol{W}\boldsymbol{\varepsilon} + \boldsymbol{u} \tag{15}$$

(b) 移動平均モデル : Moving -Average Model

$$\boldsymbol{\varepsilon} = \rho \boldsymbol{W}\boldsymbol{u} + \boldsymbol{u} \tag{16}$$

(c) Spatial Correlation Model

$$\boldsymbol{\varepsilon} = \boldsymbol{W}\boldsymbol{v} + \boldsymbol{u} \tag{17}$$

ここで、$\boldsymbol{u}, \boldsymbol{v}$ は、以下の仮定を満たす確率変数であり、$\rho$ はパラメータ、$\sigma_u : u$ の分散、$\sigma_v : v$ の分散である。

$$E[\boldsymbol{u}] = \boldsymbol{0}, E[\boldsymbol{v}] = \boldsymbol{0},$$
$$Var[\boldsymbol{u}] = \sigma_u^2 \boldsymbol{I}, Var[\boldsymbol{v}] = \sigma_v^2 \boldsymbol{I}, E[\boldsymbol{u}\boldsymbol{v}'] = \boldsymbol{0} \tag{18}$$

いずれのモデル化も、空間的自己相関の原因を空間的な相互作用に起因するものと仮定し、その相互作用の度合いを行列 $\boldsymbol{W}$ を用いて表現している。$\boldsymbol{W}$ は空間重み行列 spatial weight matrix（$I \times I$）または結合行列と呼ばれるもの

で、地点、ゾーンあるいはメッシュ間の空間的依存の度合いを表す行列である。モデルを同定可能とするためには、分析者はこの $\boldsymbol{W}$ を先験的に与える必要があり、次式のように、$\boldsymbol{W}$ の各要素 $w_{ij}$ をゾーン $i$ とゾーン $j$ の間の空間距離 $d_{ij}$ の関数として与えることが多い。

$$\begin{cases} w_{ij} = 1/d_{ij}{}^{\alpha} \;\; or \;\; \exp(-a \cdot d_{ij}) \;\; or \;\; \cdots \quad (i \neq j) & (19) \\ w_{ii} = 0 & (20) \end{cases}$$

ただし、$\alpha, a$ は定数。

パラメータの推定に最尤法を用いる場合には、$\boldsymbol{u}$, $\boldsymbol{v}$ に対して次のような正規分布が仮定される。

$$\boldsymbol{u} \sim N(0, \sigma_u^2 \boldsymbol{I}), \boldsymbol{v} \sim N(0, \sigma_v^2 \boldsymbol{I}) \qquad (21)$$

空間を対象とした場合に限らず、このようなモデルは誤差修正モデル error correction model 若しくは誤差要素モデル error component model (ECM) と呼ばれる。誤差要素モデルという用語は必ずしも現時点では一般的でないが、誤差項の要素分解 (error component) 法という用語は定着しており[6]、また、誤差修正モデルと誤差要素モデルという用語はその使われ方に違いはあるものの、本論文ではそのことは本質的な問題とならないため、以上のようなモデル化を誤差要素モデルあるいは ECM と総称する。この他にも、様々なモデル化が提案されており、それらについてのレビューは、堤他 (1998)[7] 等を参照されたい。

時系列相関に対しては、先行条件と現象との間の因果連鎖が時間の進行方向に副っており、ある時点における現象はその時点より過去へは影響を与えないという考えに基づきモデル化がなされるのに対し、空間的自己相関に対しては式(15)～(17)に示すような双方向に影響を与える構造としてモデル化されるのが特徴である。このことは、3.3 で述べるように、Kriging と整合する形で ECM に基づく空間内挿を行う際の重要な論点となる。

### 3.2 誤差要素モデルと内挿

これまでも、空間計量経済学の分野において提案されたモデルを Kriging に拡張する手法に関する研究はいくつか存在する。例えば、Cressie(1988)[8], Cressie(1993)[1]では、Markov Random Field Prediction と呼び、$\boldsymbol{\varepsilon}$ が次式に表される Markov 過程に従うとしてモデル化を行っている。

$$\boldsymbol{\varepsilon} \sim (\boldsymbol{0}, (\boldsymbol{I} - \boldsymbol{Q}^{-1})\boldsymbol{M}) \qquad (22)$$

ここで、$\boldsymbol{Q}$ は適当な形の行列であり、式(22)は計量経済学等においては、CAR(Conditional Auto Regressive)と呼ばれる定式化に他ならない。これは、共分散行列を距離のみの直接的な関数として与えており、誤差項どうしの相互作用を定式化したものではなく、この Markov Random Field Model は 2.2 で述べた Universal Kriging の 1 つの形式と考えることができる。

さらに、Haining(1988)[9], Haining(1989)[10]では同様の CAR モデルを用い、隣接する領域間の重み行列を次のような 0,1 で表される式を用いている。

$$w_{ij} = \begin{cases} 1 \\ 0 \end{cases} \qquad (23)$$

しかし、著者らがレビューを行った範囲では、3.1 に示したような空間を分析対象とした誤差要素モデルをもとに、Kriging と整合させる形で空間データの内挿を行った例は無い。

これは、一つには、ECM が主として適用される社会経済の分野では、内挿そのものを必要としない場合が多いという事実に起因するものと考えられる。例えば、人口データは市町村界等のゾーンで集計された値であり、ある地点の人口を内挿によって推定するという状況や、あるゾーンの人口データが欠如していてそれを内挿によって推定するという状況はあまり考えられない。一方、地価の分析などの際、例えば公示地価制度において、周辺の地域のある種の代表という性格を持つ点として選ばれる基準点はその数が限られているため、それ以外の場所の地価をそれをもとに推定したいという要求は大きいが、地価は一筆 (区画) ごとに、面積や形状といったその属性が大きくことなるため、通常の内挿が前提としているような変数に関する連続的な (滑らかな) 変化を仮定して内挿を行うということが、非現実的とならざるを得ない。無論、社会経済分析の対象がすべて内挿を必要としないということはなく、地価分析を例にすれば、首都圏の地価の分布図を大局的に眺めるといった場合には、一筆ずつの膨大なデータを用いた分析よりは、代表的な点のデータを用いて内挿を行うということにそれなりの意義が認められることも少なくあるまい。

空間を分析対象とした誤差要素モデルに基づき Kriging と整合する内挿手法が皆無であるもう一つの理由としては、次節で述べるように、ECM から導かれる誤差項の分散共分散行列が、結果として Kriging が前提とする定常の仮定を満足しないという問題に起因するものと考えられる。

### 3.3 誤差要素モデルに基づく Kriging による内挿の理論的検討

本節では、誤差要素モデルの考え方をもとに Kriging を用いて空間内挿を行うことが可能であるか、それが可能でないとするとどのような問題によるものであるかついて考察を行う。

回帰モデルの誤差項に対し、3.1 に示した(a)～(c)のようなモデルを導入した場合、その分散共分散行列は次式のとおり与えられる。

$$\text{(a)} \quad Cov(\varepsilon) = \sigma_u{}^2 \left[ (\boldsymbol{I} - \rho \boldsymbol{W})' (\boldsymbol{I} - \rho \boldsymbol{W}) \right]^{-1} \qquad (24)$$

(b) $Cov(\varepsilon) = \sigma_u^2\left[(I+\rho W)'(I+\rho W)\right]$ (25)

(c) $Cov(\varepsilon) = \sigma_v^2 W'W + \sigma_u^2 I$ (26)

時間座標と空間的な座標を同一視すれば、時系列相関と空間的自己相関は本質的には同じであるとも考えられる。しかし、時系列相関に対しては、先行条件と現象との間の因果連鎖 が時間の進行方向に副っており、ある時点におけるの現象はその時点より過去へは影響を与えないという考えに基づきモデル化がなされる。これに対し空間的自己相関の場合には、3.2 で述べたように、双方向に影響を与える構造としてモデル化されることが一般的であり、時系列とは根本的に異なる側面を持つ。（ただし、後者についても時間軸を導入した時空間モデルへと拡張し、時間軸上で一方向へ影響を与える構造とすることも可能である。）この空間特有の双方向の影響、すなわちフィードバックが、時系列モデルと本質的に異なる結果をもたらす。このことは、誤差項の（自己）分散（共分散行列の対角項）について考察すると明確になる。何故なら、地点 $i$ における自己分散は式(27)に示されているとおり、自分自身の分散の他に他のあらゆる地点を介した影響を受けている（図２）。

$$Var(\varepsilon_i) = \sigma_u^2\left[1+\rho^2\sum_{k=1}^{n} w_{ik}w_{ki}\right] \quad (27)$$

従って、これらの地点がある特殊な配置に分布している場合を除いて、空間を対象とした AR あるいは MA モデルによって導かれる誤差項の分散共分散行列は、意図するか否かに関わらず、$\rho>0$ では結果として不均一分散を生じさせることとなる。

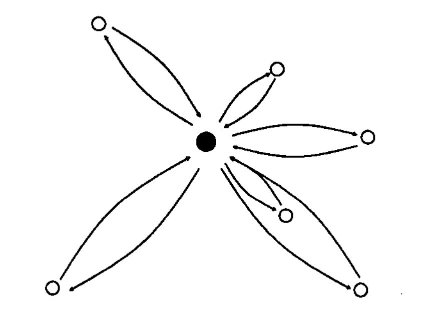


**図２．自己分散に影響を及ぼす空間相互作用**

一方、任意の異なる２地点 $i, j$ 間における共分散、例えば式(25)に示す誤差項の共分散行列の各非対角要素は次式のように表わされる。

$$Cov(\varepsilon_i,\varepsilon_j) = \sigma_u^2\left[\rho(w_{ij}+w_{ji})+\rho^2\sum_{k=1}^{n} w_{ik}w_{kj}\right] \ (i \neq j) \quad (28)$$

この式が示すとおり、異なる２地点 $i, j$ 間における共分散は、それぞれの地点と別の地点 $k'$（$\neq i,j$） との相関によって影響されるため、やはり、これらの地点がある特殊な配置に分布している場合を除いて、２地点 $i,j$ 間の距離 $d_{ij}$ のみの関数とはならない（図３）。この結果、地点 $i$ から等距離にある地点 $j$ と地点 $k$ との間の共分散が異なるという結果が生じる。

このように、一般に、ECM によってモデル化された誤差項の共分散は、Kriging で前提とされる定常性を満足しないことが分かる。これらの点は、時系列モデルにはない大きな特徴であると言える。実際、時系列モデルの内挿については、例えば時系列の AR モデルに関して Kriging と整合する内挿手法に関する研究が比較的早い時期に Arthur(1962)[11]によって行われている。

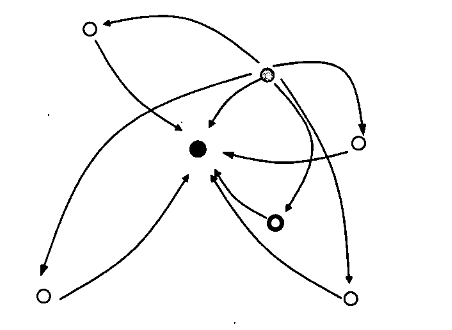


**図３．共分散に影響を及ぼす空間相互作用**

## 4. 誤差要素モデルに基づく Kriging による空間内挿法の提案

3.で述べたように、ECM に基づき Kriging と整合する空間内挿を行うに際しては、結果的に誤差項の分散共分散行列が定常性の仮定から導かれる性質を満足しないという問題が生じる。そこで本章では、ECM のフレームを維持したまま、誤差項の分散共分散が距離のみの関数となるような方法を提案する。

まず、図４に示すように、実際の観測地点を含む地点が、無限に広がる格子点状に配置されており、観測されていない地点はそのうちのデータが欠損している地点であると想定する。

この時、格子点の数が無限のまま誤差要素モデルを適用すると問題が生じる。何故なら、3.3 で説明したように、地点 $i$ の自己分散は他の点の影響を受け、また、地点 $i$ と地点 $j$ の共分散も同様にこれら以外の点の影響を受けるため、無限に多くの点を考慮することによって発散して

しまうからである。そこで、ある地点への影響は、ある有限の領域内のみに限るといった仮定を設けることとする。

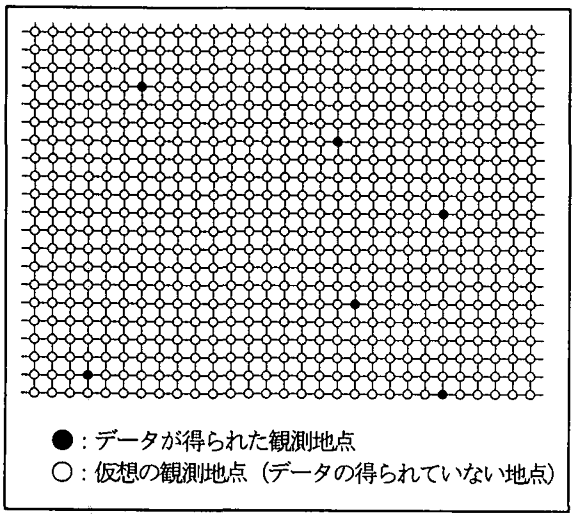


**図4．仮想観測地点の設定**

この際、各点に関しての自己分散の計算において、影響を考慮する地点数を同一にしなければ不均一分散をもたらすため、ある一定の範囲、具体的には半径 $R$ の円内に含まれる地点のみを考慮することとする（図5）。

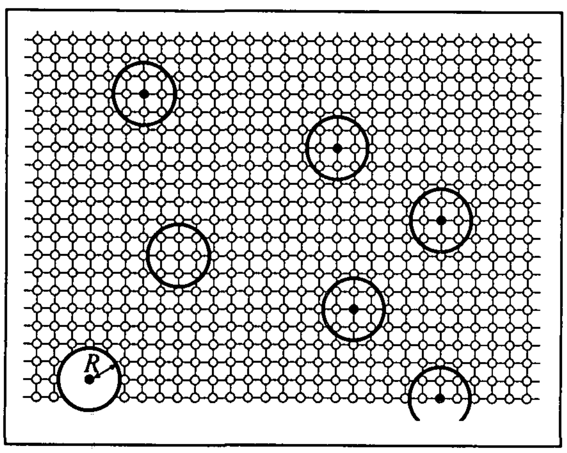


**図5．range の設定**

すなわち、重み $w_{ij}$ に対し、式(29)のような制約を置く。

$$w_{ij} = \begin{cases} f(d_{ij}) & (0 \le d_{ij} \le R) \\ 0 & (R < d_{ij}) \end{cases} \tag{29}$$

このような空間的な制約は、Kriging において range と呼ばれるものに相当する。共分散についても、range を設定することで項が発散しないような構造となる。range を導入すると、結局のところ格子に無限を仮定する必要はなく、観測点が分布する領域から range 程度分だけ外に広がった領域内において格子を仮定すれば良いことになる。

このようにして共分散関数を定義することにより、(b)の MA モデルに関しては、その考え方に基づき Kriging の基本的な前提と整合する内挿が可能となることが確認できる。

このことを、次のような簡単な場合について説明する。図6に示すように、格子の間隔を $\bar{d}$ とする。さらに、range の半径 $R$ は次の式(30)を満たす定数として、重みは式(31)の関数によって与えられるものとする。

$$\sqrt{2}\bar{d} \le R < 2\bar{d} \tag{30}$$

$$w_{ij} = 1/d_{ij}{}^{\alpha} \quad (i \ne j) \ , \ w_{ii} = 0 \quad (\alpha : 定数) \tag{31}$$

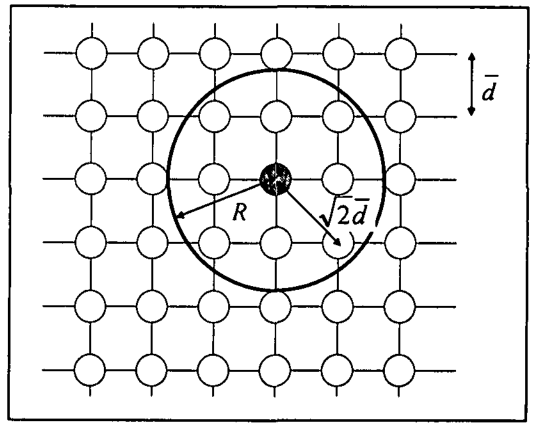


**図6．自己分散に影響を及ぼす範囲：格子間隔 $\bar{d}$，range の半径 $\sqrt{2}\bar{d} \le R < 2\bar{d}$**

この時、式(27)から自己分散が次式のように導かれる。

$$Var(\varepsilon_i) = \sigma_u{}^2\left[1 + \rho^2\left\{4\cdot\left(\frac{1}{\bar{d}^{\alpha}}\right)^2 + 4\cdot\left(\frac{1}{\sqrt{2}\bar{d}^{\alpha}}\right)^2\right\}\right]$$
$$= \sigma_u{}^2\left(1 + 6\rho^2\frac{1}{\bar{d}^{2\alpha}}\right) \tag{32}$$

式(32)は、自己分散がどの点においても等しいことを示している。

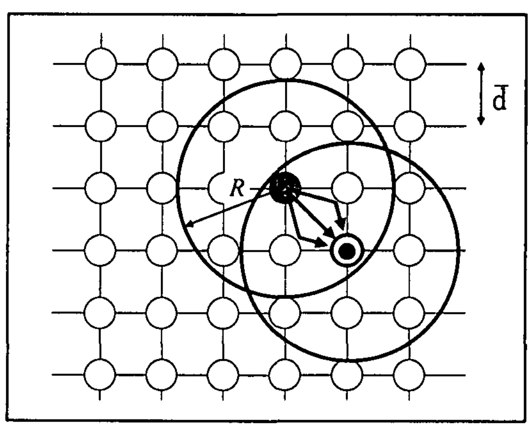


**図7．共分散に影響を及ぼす範囲：格子間隔 $\bar{d}$，range の半径 $\sqrt{2}\bar{d} \le R < 2\bar{d}$， $0 < d_{ij} \le R$**

一方、共分散については、例えば距離が$\sqrt{2}\bar{d}$ だけ離れた隣接する２点の共分散は次式のように導かれる。

$$Cov\left(\varepsilon_i,\varepsilon_j\right)=\sigma_u^2\left[\rho\left(2\cdot\frac{1}{\bar{d}^{\alpha}}\right)+\rho^2\left\{2\cdot\left(\frac{1}{\bar{d}^{\alpha}}\right)^2\right\}\right] \tag{33}$$

これらの式(32)･(33)のいずれにおいても、右辺は格子の間隔$\bar{d}$ のみの関数となっている。すなわち、2.1 において説明した定常の仮定を満足している。

なお、異なる２点 $i,j$ における値 $\varepsilon_i, \varepsilon_j$ の共分散については、以下に示すように、range の半径 $R$ と２点間の距離 $d_{ij}$ との関係に応じて、それぞれ関数形が異なる。

(i) $0<d_{ij}\le R$　　(ii) $R<d_{ij}\le 2R$　　(iii) $2R<d_{ij}$

これは、式(29)によって、ある点の影響範囲が range の半径 $R$ 以内に制約されていることによるものであり、具体的には、以下のような理由による。

- (i)においては、$i, j$ 以外の点を介した作用に加え、$i, j$ の直接的な作用が共分散に影響を及ぼす（図７参照）。
- (ii)においては、２点 $i,j$ 間の距離が range の半径より大きいため、この２点の直接的な作用は共分散に影響を及ばさず、$i, j$ 以外の点を介した作用のみが共分散に影響を及ぼす。
- さらに(iii)においては、$i,j$ 以外の点を介した作用も生じないために共分散が 0 となる。

これらについての詳細な定式化は、本論文では省略することとする。

既述のように、実際には内挿を行う点の数は有限個で良いため、無限に代えて十分多くの数の格子を想定する。格子点の数を $N\times N$ とすると、対応する空間重み行列結合 $W^{(N)}(N\times N)$は式(34)のように定義される。

$$W^{(N)}=\begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{pmatrix} \tag{34}$$

ここで、観測データの入手が可能なのは図４において●で表された $I$ 地点のみであり、観測データより導かれる回帰式は以下のとおり $I$ 個となる。

$$y=X\beta+(I+\rho W^{(I)})u \tag{35}$$

ただし、$W^{(I)}$ は $I\times N$ の行列であり、以下に示すように $W^{(N)}$ から、仮想的に設定されたためにデータの得られていない格子点に対応する行を削除した行列である。

$$W^{(N)}=\begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ \cancel{w_{21}} & \cancel{w_{22}} & \cancel{\cdots} & \cancel{w_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \cancel{w_{N1}} & \cancel{w_{N2}} & \cancel{\cdots} & \cancel{w_{NN}} \end{pmatrix}$$

$$\Rightarrow W^{(I)}=\begin{pmatrix} w_{11} & w_{12} & \cdots w_{1I} \\ \vdots & & \ddots \ \vdots \\ w_{I1} & w_{I2} & \cdots w_{II} \end{pmatrix} \tag{36}$$

このとき、次式(37)の誤差項に関する分散共分散行列は式(38)のとおり与えられる。

$$\varepsilon=(I+\rho W^{(I)})u \tag{37}$$

$$C=\sigma_u^2(I+\rho W^{(I)})^t(I+\rho W^{(I)}) \tag{38}$$

これより、通常の ECM 同様、最尤推定法により各パラメータが推定される。

## 5. おわりに

本論文では、空間計量経済学においてよく用いられてきた ECM に、空間統計学における共分散関数の概念を取り込むことにより、ECM の概念をもとに空間内挿が行える手法を提案した。具体的成果は以下のとおりである。

（１）　計量経済学の分野で発展してきた従来の誤差要素モデル(ECM)の空間概念の延長上では、空間統計学の分野における推定手法である Kriging 手法をそのまま用いた空間内挿を行うことはできないことを示した。その理由は、ECM では結果として誤差項の分散不均一を生じさせ、また、任意の２地点間の共分散がその地点間の距離のみの関数にならないということである。すなわち、ECM における誤差項は、Kriging において必要となる確率変数の定常性を満足しないためである。

（２）　これらの問題点を解決するため、観測地点と内挿を行う地点を含む無数の地点が正方格子状に配置されたシステムを導入した。その無限格子において、観測データが得られていない点はデータが欠損しているものと想定することにより、ECM の一手法である移動平均（MA）モデルに基づき、通常の Kriging と整合する形での空間内挿が可能となる方法を提示した。

今後の研究課題としては、実際のデータを用いて本研究で提案した手法と既存の Kriging を用いた内挿を行い、実証比較を行うことがあげられる。これについては、現在、降雨のデータを検討に着手したばかりであり、別の機会に改めて結果を発表したい。

また、本研究では移動平均モデルについてのみ、拡張の方向性を提示した。自己回帰（AR）モデルでは、Kriging と整合する形でのモデルの拡張が不可能であると予想されているが、Spatial Correlation Model に関しては、まだ十分な検討を行っていない。これらについては、今後の課題としたい。

## 謝辞

匿名の査読者から、本稿に関し非常に有益な助言を賜った。ここに記して深く感謝の意を表す。


## 参考文献


1) Cressie, N.: *Statistics for Spatial Data, Revised Edition*, John Wiley & Sons, 1993.
2) Bailey, T. C. and Gatrell, A. C.: *Interactive Spatial Data Analysis*, Longman, 1995.
3) 間瀬 茂：『Spatial Statistics 空間統計学（空間点過程理論とその応用）』，慶應義塾大学，Seminar on Mathematical Sciences，No.24，1997.
4) Anselin, L.: *Spatial Econometrics: Methods and Models*, Kluwer Academic, 1988.
5) Kelejian, H. and Robinson, D. : A Suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model, *Papers in Regional Science*, Vol.72, pp.297-312, 1993.
6) 竹内啓編集代表：『統計学辞典』，東洋経済新報社, 1989.
7) 堤盛人・清水英範・福本潤也・井出裕史： 誤差項に空間的自己相関が存在する回帰モデルのパラメータ推定手法に関する考察，土木計画学研究・論文集, No.15，pp.49-56，1998.
8) Cressie, N.: Spatial Prediction and Ordinary Kriging, *Mathematical Geology*, Vol..20, No.4, pp.405-421, 1988.
9) Haining, R.: Estimating Spatial Means with an Application to Remotely Sensed Data, *Communication in Statistics A - Theory and Methods*, Vol.17, No.2, pp.573-597, 1988.
10) Haining, R.: Maximum Likelihood Estimation with Missing Spatial Data and with an Application to Remotely Senced Data, *Communication in Statistics A - Theory and Methods*, Vol.18, No.5, pp.1875-1894, 1989.
11) Arthur, S.G.: Best Linear Unbiased Prediction in the Generalized Linear Regression Model, *Journal of American Statistical Association*, Vol.57, pp.369-375, 1962.



(2000 年 4 月 21 日 受付)