【原著論文】

ネットワークを介した空間的相互作用を考慮した
面域における空間内挿

村上大輔*・堤盛人**

Spatial prediction of data in a plane considering the spatial interaction enforced by a network

Daisuke MURAKAMI*, Morito TSUTSUMI**

Abstract: In traditional spatial statistics, the spatial dependence between two arbitrary sites on a plane is expressed in terms of the Euclidean distance between the two sites. However, in actual cities or regions, infrastructure networks such as railway networks cause spatial interaction and thus strongly affect the spatial dependence. We propose a new spatial prediction method, based on a primal spatial statistical method called Kriging, that explicitly considers such dependence. We first divide the entire plane into Voronoi regions and configure a network node in each region to create a hypothetical network. Then, in order to investigate the spatial interaction enforced by the network, we estimate the shortest-path distance by considering a transportation network and present the covariance function of that distance. Next, we combine the covariance functions based on both the shortest-path distance and the Euclidean distance. Finally, we assess the effectiveness of the proposed method by applying it to a case study of land price prediction.

Keywords: 空間従属性 (spatial dependency), 空間的相互作用 (spatial interaction), ネットワーク (network), クリギング (Kriging), 最短経路距離 (shortest path distance)

1. はじめに

一般に、空間的に近接したデータ間には類似した性質が見られることが知られている。この性質は空間従属性と呼ばれ、空間データ分析の際に考慮すべき重要な性質の一つである。

空間従属性を考慮した研究の盛んな学問分野の一つに空間統計学がある。空間統計学は自然科学に発展経緯を持ち、従来、共分散行列を共分散関数と呼ばれるユークリッド距離の関数に基づいて与えることで空間従属性が考慮されてきた。しかしながら、従来の共分散関数では空間従属性を捉えきれないと示唆されることも多く、より柔軟に空間従属性を考慮するための様々なアプローチが試みられてきた。例えば、柔軟な関数に基づく共分散関数の構築 (e.g. Diggle and Verbyla, 1998), 補助データとの相互

共分散を考慮した空間従属性の定式化 (e.g. Cressie, 1993), ユークリッド距離に代わる距離指標の適用 (e.g. Curriero, 2006) などが試みられてきた。近年、3つ目に挙げた、ユークリッド距離に代わる距離指標に基づく共分散関数構築の研究が注目されており (Gotway and Young, 2008), 様々な距離に基づく共分散関数の構築が試みられつつある。

本研究も、ユークリッド距離に代わる距離指標を適用した空間として最短経路距離に着目する。

近年、空間統計学の社会経済データへの適用が、盛んに試みられており、その有用性が多くの研究により示されている。しかしながら、それらの研究の殆どは、未だにユークリッド距離に基づく共分散関数をそのまま用いており、社会経済データ特有の空間従属性の特性を見落としている可能性がある。

* 学生会員 筑波大学大学院システム情報工学研究科 (University of Tsukuba)
〒305-8573 つくば市天王台 1-1-1
E-mail : muraka51@sk.tsukuba.ac.jp

** 正会員 筑波大学大学院システム情報工学研究科 (University of Tsukuba)

そこで本研究は、空間統計学における最も代表的な空間内挿手法である Kriging (e.g. Cressie, 1993) を社会経済データへ適用することを念頭に、社会経済活動に対して特に大きな影響を及ぼす交通ネットワークに着目し、交通ネットワークに起因した空間的相互作用が引き起こす空間従属性を明示的に考慮する方法について提案する。

以降では、第2章で Kriging の理論を概観した後、第3章において本研究に関連する既存の研究を整理し、本研究の新規性について説明する。そして第4章において、上述の目的に則し、最短経路距離に着目した共分散関数の提案を行う。最後に第5章において、提案した共分散関数を Kriging に導入して実際のデータに適用することで、その有用性を検証する。

2. Kriging の概要

空間統計学の代表的な手法である Kriging は、ユークリッド空間上の地点 $s_i \in R^2 (i=1, \dots, n)$ で観測された被説明変数を用いて、地点 $s_0 \in R^2$ の被説明変数 y_0 を予測することのできる手法であり、(1) 式を基本式とする。

$$y = X\beta + \varepsilon \quad (1)$$

ここで、 y は $n \times 1$ の被説明変数ベクトル、 X は $n \times p$ の説明変数行列、 β は $p \times 1$ のパラメータベクトルであり、 $X\beta$ はトレンドを表す。また、 ε は $n \times 1$ のトレンドでは説明されない局所の変動を表し、(2)式、(3)式で定義される二次定常性に従うと仮定する。

$$E[\varepsilon_i] = 0 \quad (2)$$

$$\text{Cov}[\varepsilon_i, \varepsilon_j] = C(d_{ij}) \quad (3)$$

d_{ij} は距離であり、通常、ユークリッド距離が仮定される。 $C(d_{ij})$ は共分散関数であり、これにより空間従属性を表現する。共分散関数には、例えば指数型共分散関数(4)式が存在する。

$$C(d_{ij}) = \begin{cases} \sigma^2 \exp\left[-\left(\frac{d_{ij}}{w}\right)\right] & (s_i \neq s_j) \\ \tau^2 + \sigma^2 & \text{otherwise} \end{cases} \quad (4)$$

ここで、 σ^2 は partial-sill、 τ^2 は nugget、 w は range と呼ばれるパラメータである (Cressie, 1993)。

基本式(1)を踏まえ、予測地点 s_0 の被説明変数 y_0 は(5)式に従うと仮定する。

$$y_0 = x_0' \beta + \varepsilon_0 \quad (5)$$

ここで、 x_0 は地点 s_0 の $p \times 1$ の説明変数ベクトル、 ε_0 は局所の変動(スカラー)を表す。 y_0 の予測は、期待二乗誤差分散の最小化(6)式に基づき行われる。

$$\min [Var(y_0 - \hat{y}_0)] \quad (6)$$

期待二乗誤差分散の最小化の結果として、 y_0 の予測量は(7)式で与えられることが知られている。

$$\begin{aligned} \hat{y}_0 &= x_0' \hat{\beta} + c_0' C^{-1} (y - X\beta) \\ \hat{\beta} &= (XC^{-1}X)^{-1} XC^{-1}y \end{aligned} \quad (7)$$

ここで、 C は観測地点間の共分散行列、 c_0 は観測地点と予測地点との間の共分散ベクトルをそれぞれ表し、サイズはそれぞれ $n \times n$ 及び $n \times 1$ である。(7)式で与えられる Kriging の予測量は、最良線形不偏予測量 (BLUP: Best Linear Unbiased Predictor) という、統計的に優れた性質を持つ。

Kriging のためのパラメータ推定手法のうち、本研究では、(8)式の最大化に基づいてパラメータを推定する最尤法を用いることとする (Cressie, 1993)。

$$L = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln(y - X\beta)' C^{-1} (y - X\beta) - \frac{1}{2} \ln|C| \quad (8)$$

3. 関連する既存研究と本研究の位置付け

社会科学を中心に空間従属性を考慮した研究が進められてきた空間計量経済学では、空間的相互作用が空間従属性の強い誘因であるとされる (Anselin, 1988)。従来、空間計量経済学においても、空間従属性は地点(ゾーン)間の単純なユークリッド距離を用いて定式化されてきたが、近年、多くの研究で、空間的相互作用の強い誘因と示唆される社会経済距離が、空間従属性の定式化に用いられるようになってきている。例えば、Conley and Giorgio (2000) は、失業率の空間従属性の定式化に複数の社会経済距離の適用を試み、Dall'erba (2005) は労働生産性の空間従属性に時間距離の適用を試みている。

一方、空間統計学においても、社会経済データへの応用が盛んに試みられるようになり、その有用性

が多くの研究により示されている (e.g. Tsutsumi and Seya, 2009). その空間統計学では、従来から、空間従属性が方位に依存していることに注目した異方的共分散関数 (Anisotropic covariance function) (e.g. Zimmerman, 1993) など、単純なユークリッド距離では捉えきれない空間従属性を表現する方法が提案されている。Gotway and Young (2008) は、非ユークリッド距離を考慮するための空間統計学のアプローチを、等長埋め込み (Isometric embedding) に基づくアプローチ (e.g. Curriero, 2006) と、多次元尺度構成法 (Multidimensional scaling) に基づくアプローチ (e.g. Barry and Ver Hoef, 1996) に分類し、その上で、座標変換に基づく手法である前者では、座標変換の際に必然的に誤差が発生してしまうのに対し、後者では誤差の発生なしに非ユークリッド距離を考慮できることを指摘している。

多次元尺度構成法に基づく代表的なアプローチとして、データ間の類似の尺度として用いられることの多いカーネルを、構造を持つ空間上に分布するデータ間の類似の尺度に拡張した畳み込みカーネル (Convolution kernel) を用いるアプローチがある。Barry and Ver Hoef (1996) は、畳み込みカーネルを用いた、空間従属性の柔軟な定式化のための方法論を提示しており、彼らの方法論を応用する形で、非ユークリッド距離に基づく共分散関数の研究が行われつつある。例えば、Ver Hoef et al. (2006), Cressie et al. (2006), Garreta et al. (2010) は、河川ネットワーク上の最短経路距離に基づく空間従属性の考慮を行うことのできる手法を提案した。しかしながら、それらの手法は、河川ネットワークがツリー構造、即ち、任意地点間の最短経路が一通りのみのグラフ構造であることを前提に議論を行っているため、そのまま一般のネットワークに適用することはできない (詳しくは次章 3.2 節)。さらに、それらは線分から構成される (局所的に 1 次元の) 河川ネットワーク空間内を対象とした内挿のための手法であり、河川ネットワーク外の空間は、明示的には考慮されない。実際、これらの研究の適用事例においても、例えば、4.4 節で詳述するように、Cressie et al. (2006) では河川ネットワーク上の距離に加えて、距離指標とし

てユークリッド距離も考慮しており、暗黙に河川ネットワーク空間の外の空間の存在も仮定されているものの、予測の対象とする空間はあくまでも (局所的に 1 次元の) 河川ネットワーク内の空間に限られている。従って、本研究で対象とするような、例えば鉄道ネットワークの存在を考慮した都市空間全体、即ち、鉄道ネットワークの外の空間での予測 (内挿) にそれらを援用するためには、ネットワーク外の空間を予測の対象とするための理論的拡張が必要となる。

2 次元の平面上で最短経路距離を考慮した空間統計学の手法としては、筆者らの知る限り、Krivoruchko and Gribov (2004) が唯一の研究である。彼らの方法は、対象地域を細かなグリッドで分割し、隣接するグリッド間に与えられたコスト距離の足し合わせにより与えられる任意グリッド間の距離を用いて空間従属性を定式化するという素朴なものである。当然のことながら、グリッド単位でしか内挿ができないため、平面上の任意の地点の空間内挿ができず、交通ネットワークの考慮は大雑把なものにならざるを得ない。

以上のことを踏まえた上で、本研究では、都市・都市圏の構造を規定する最も基本的な社会経済距離の一つである最短経路距離に基づいて共分散関数を構築し、交通ネットワークに起因した空間的相互作用をより明示的に考慮し、都市・都市圏内の任意の地点で特に社会経済データの空間内挿が可能な手法を、空間統計学の理論に整合する形で提示する。これ以降、対象とする空間全体の巨視的な構造とそこでの空間的相互作用に大きな影響を及ぼす交通ネットワークとして、駅 (ノード) を持つ鉄道ネットワークを考える。ただし、実際には、航空や高速道路などの交通ネットワークに適用可能である。一方、駅周辺の微視的な構造は、駅とこれを中心とした道路網によって形成されるものとする。詳しくは、4 章 3 節において説明する。

4. 最短経路距離を考慮した共分散関数の提案

4.1. 提案の前提

本章では、鉄道ネットワークを通じた空間的な相

相互作用が引き起こす空間従属性を、最短経路距離に基づく共分散関数の定式化により、表現する。しかしながら、多くの空間統計学の既存研究が仮定するように、ユークリッド距離もまた、共分散関数に考慮することが自然である。ユークリッド距離を用いることで、自然環境の空間的な連続性など、多様な空間従属性も考慮できる。従って、以下では、最短経路距離及びユークリッド距離の両方を考慮した共分散関数を提案する。

4. 2. 距離及び共分散が満たすべき性質

非ユークリッド距離に基づく共分散関数の構築に際しては、構築された関数の理論との整合性を示さなくてはならない (Gotway and Young, 2008). Curriero (2006) はユークリッド距離の関数として定義された従来の共分散関数に非ユークリッド距離を投入した場合、共分散としての十分条件である正定値性 (後述) を満足しなくなることを例示している。

ここで、共分散の十分条件は以下で定義される。

$$\text{Cov}(\varepsilon_i, \varepsilon_i) \geq 0 \quad (9-a)$$

$$\text{Cov}(\varepsilon_i, \varepsilon_j) \geq \text{Cov}(\varepsilon_j, \varepsilon_i) \quad (9-b)$$

$$\sum_i^n \sum_j^n k_i k_j \text{Cov}(s_i, s_j) \geq 0 \quad k_i, k_j = \text{const.} \quad (9-c)$$

(9-c) は半正定値性と呼ばれ、分散が非負であるための必要十分条件である。しかしながら、Kriging の予測式、及び、尤度関数には共分散行列の逆行列が含まれており、これらが解を持つためには、半正定値性よりも強い条件である正定値性、即ち、(9-c) から等号を除いた次式を満足しなくてはならない。

$$\sum_i^n \sum_j^n k_i k_j \text{Cov}(s_i, s_j) > 0 \quad k_i, k_j = \text{const.} \quad (9-c')$$

従って、以下では、(9-a)、(9-b) 及び (9-c') を必ず満足する共分散関数の導出を行う。

一方で、共分散関数は、以下の条件を満足する距離に基づいて空間従属性を考慮するものとして定義される (Krivoruchko and Gribov, 2004)。

$$d_{ij} \geq 0, \text{ かつ } i=j \text{ の場合のみ } d_{ij}=0 \quad (10-a)$$

$$d_{ij}=d_{ji} \quad (10-b)$$

$$d_{ij} \leq d_{ik} + d_{kj} \quad (\text{三角不等式}) \quad (10-c)$$

(10-a) ~ (10-c) は距離の十分条件として知られ、これらを満足することは、距離に基づいて与えられる空間の連続性を満足するための必要十分条件として知られている。従って、上記の条件を満足しない距離指標を連続的な空間を前提とする空間統計学で用いることはできない。

従って、以降では、距離の十分条件 (10-a), (10-b), (10-c) を満足する最短経路距離に基づいて、共分散の十分条件 (9-a), (9-b) 及び (9-c') を満足する共分散関数を構築し (4 章 3 節)、ユークリッド距離に基づく従来の共分散関数と結合することで両距離を考慮した共分散関数を構築する (4 章 4 節) こととする。

4. 3. 最短経路距離に基づく共分散関数

4. 3. 1. 提案の前提

鉄道をはじめとする公共交通は、通勤に代表される大量輸送を担った結果として、基幹交通路を形成する。その一方で、公共交通へのアクセス向上のために、駅を中心として交通システムや道路ネットワークが整備される (竹内, 2000)。以上のような鉄道整備や集客の結果として駅が存在が影響を及ぼす圏域は駅勢圏と呼ばれる。これらのことを念頭に、本研究では、以下のような構造を持つ都市あるいは都市圏を想定する。

[1] 対象とする空間は、通勤鉄道ネットワークによって形成された駅勢圏の集合により稠密に覆われており、駅勢圏間の全トリップは鉄道を介して行われる。鉄道を介する全トリップの起点は乗車駅の駅勢圏内に、終点は下車駅の駅勢圏内に存在する。

[2] 各駅勢圏の中は駅を中心とした都市 (街) 構造を持ち、道路ネットワークによってその構造が支配される。

上述の [1]・[2] の結果として以下のような空間的相互作用によって空間従属性が生まれると考える。

[1] 対象とする空間における、鉄道ネットワークを介した駅勢圏間の巨視的な空間的相互作用。

[2] 各駅勢圏内における、道路ネットワークを介し

図1 本研究での都市 / 都市圏の捉え方と空間従属性のイメージ

た任意地点間の微視的な空間的相互作用。

これらの空間的相互作用の結果として生じる空間従属性は、図1に概念的に示すように、鉄道ネットワーク上、及び、各駅勢圏上を伝播しながら、距離の増加に伴い次第に逓減していくものとする。各駅 $\{I, J, K, \dots\}$ は、駅勢圏 $\{D_I, D_J, D_K, \dots\}$ をそれぞれ形成し、 D_I 内の任意地点は $\{s_i, s_j, s_k, \dots\}$ であるとする。

4.3.2. 共分散関数の定式化

本研究では、前節の [1] 及び [2] に起因した空間従属性を考慮することのできる共分散関数の構築を、下の [1] 及び [2] でそれぞれ行い、下の [3] で両者を結合することで最短経路距離を考慮した共分散関数を構築する。

[1] 鉄道ネットワークによって規定される駅勢圏間の共分散関数の構築

[2] 各駅勢圏内の地点間の共分散関数の構築

[3] [1]・[2] で構築された両共分散関数の結合による、任意地点間の共分散関数の構築

以下、[1], [2], [3] の順で詳述する。

[1] 鉄道ネットワーク上の空間従属性を説明しうる距離指標として、以下を用いることとする。

dn : 鉄道ネットワーク上のノード間の最短経路距離

dn は明らかに距離の十分条件 (10-a), (10-b) 及び (10-c) を満足する。

次に、 dn に基づく共分散関数を定義する。ここでは、河川ネットワーク上の最短経路距離に基づいて共分散関数を定義した研究の一つである Garreta et al. (2010) を援用する。彼らは、河川ネットワークの構造がツリー、即ち、任意地点間の最短経路が

一通りのみのグラフであることを前提に、ツリー上の最短経路距離に基づく共分散関数を構築した。その際、共分散の十分条件を満足する共分散関数の構築のために、畳み込みカーネルを援用することで、ツリー上の地点 s_0 の局所変動成分 ϵ_0 (詳しくは2章) を (11) 式で定義した。

$$\epsilon_0 = \int_{-\infty}^{\infty} g(s-s_0|\boldsymbol{\theta}) u(s) ds \quad (11)$$

ここで、 $u(s)$ は1次元ユークリッド空間 (ツリーの各リンクに相当) 上の任意地点 s で観測されるランダムノイズ、 $g(s-s_0|\boldsymbol{\theta})$ は s_0 から発生させたカーネル、 $\boldsymbol{\theta}$ はパラメータベクトルを表す。(11) 式で定義される ϵ_0 の共分散を算出した結果として導出することのできる関数は、共分散の十分条件 (9-a), (9-b) 及び (9-c') を満足することが知られている (e.g. Ver Hoef et al., 2006)。

Garreta et al. (2010) では、ツリー上の最短経路距離の関数として定義された指数型共分散関数 (4) 式が、(11) 式から導出可能であることを示すことで、ツリー上の最短経路距離で定義された指数型共分散関数が、共分散の十分条件を満足することを示した。具体的には、ツリー上の2地点の配置パターンを、一方が他方の上流 (下流) にある場合とそうでない場合とに分類し、(11) 式を用いた指数型共分散関数の導出が可能であることを各々について示した。しかしながら、交通ネットワークのような一般のネットワーク上の地点は、ツリーの場合と同様の配置パターンの分類ができないため、指数型共分散関数の適用性の確認が必要である。

そこで、指数共分散関数の一般のネットワークに対する適用性を、Garreta et al. (2010) における知見と拡張最短距離木 (例えば四茂野, 1993) の概念を用いることで確認する。拡張最短距離木とは、ネットワーク上の任意地点を基点とし、そこからの最短経路に基づいてツリーに変換したグラフである。例えば、図2-(a)で示されるネットワーク上の地点 s_p からの拡張最短距離木は図2-(c)であり、 s_p からの最短経路距離は、 s_p を基点とした拡張最短距離木上の最短経路距離と等しい。従って、地点 s_p と任意地点との間の最短経路距離は、ツリー上の距離とし

図2 拡張最短距離木に基づくネットワークの変換

て定義されるため、同距離で定義された指数型共分散関数は共分散の十分条件を満足する。同様に、ネットワーク上のどの地点を基点としても、拡張最短距離木上の距離として定義することができる。従って、図2-(a)で示されるネットワーク上の任意の地点間の最短経路距離で定義された指数型共分散関数は共分散の十分条件を満足する。また、同様の議論は、明らかに、一般のネットワークに対しても適用可能である。

以上より、鉄道ネットワーク上の距離指標として dn を導出し、 dn に基づく共分散関数として指数型共分散関数が適用可能であることを明らかとなった。

[2] 駅勢圏内の任意地点の間の空間的相互作用は、駅周辺の交通ネットワークに強く依存すると考えるのが自然であろう。しかしながら、実際の都市における道路ネットワークは複雑であり、さらに道路を用いるのは車だけでなく徒歩やバスなど交通手段があるため、実際の適用を考えた場合には、それらを同時に考慮するのはデータ入手の面からあまり現実的とは言えまい。そこで本研究では、個別の駅勢圏内の具体的な交通ネットワークについては情報を持たないことを前提に共分散関数を定義する。

竹内（2000）の指摘にもあるように、どのような駅においても駅勢圏内からの一定の交通の集中が見られることを考えると、各駅は各駅勢圏内の核であると考えるのは自然であろう。そこで本研究では、各駅勢圏内の領域に関して以下のような仮定を設ける。

①駅から放射状に延びる交通路に沿って性質の似た

地域が伸びる

②似た性質を持つ地域が駅からの環状方向に広がる

その上で、①、②のような現象を通した空間従属性を考えることとする。具体的には、各駅勢圏に核である駅を中心とした稠密な仮定の放射環状道路を想定する（図3）。そして、同一駅勢圏内の2点 i と i' について、①を表現するための距離指標として dr を、②を表現するための距離指標として dn をそれぞれ定義する。

$dr_{i,i'}$ ：最寄り駅までの直線（放射）距離の差

$da_{i,i'}$ ：2点 i と i' を通る環状道路のうち、駅に近い方の道路に沿った円弧の長さの差

その上で、 dr 、 da の距離抵抗の違いを考慮した、距離指標（12）式を定義する。

$$d_{i,i'}^* = \frac{dr_{i,i'}}{w_{dr}} + \frac{da_{i,i'}}{w_{da}} \quad (12)$$

w_{dr} と w_{da} は、 dr 及び da の距離抵抗を表すパラメータである。ここで定義した距離指標(12)式は、 $w_{dr} = w_{da} = 1$ の時、放射環状距離（いわゆる Karlsruhe 距離）となる。本研究では、次章の実証分析で社会経済データとして用いる地価の分布などの例では、上述のように放射方向と環状方向で空間従属性が異なると考え、このような距離指標を用いる。

距離指標(12)式は距離の十分条件を満足しなくてはならない。距離指標(12)が条件(10-a)、(10-b)を満足することは明らかである。しかしながら、(12)式が、必ずしも条件(10-c)を満足するとは限らない。例えば、ノード I を挟んで対極に地点 s_i 及び地点 $s_{i'}$ が位置する場合、 s_i 、 I 間の距離および I 、 $s_{i'}$ 間の距離の和は、 s_i 、 $s_{i'}$ 間の距離よりも短くなる。そこで、

図3 駅勢圏内の距離指標の定義

(12)式で定義される距離指標への制約として(10-c)を与えると、(10-c)を満足するための制約式(13)が導出される。

$$w_{dr} \leq (2/\pi)w_{da} \quad (13)$$

従って、(12)式で定義される距離指標の制約として、(13)式を与えることとする。

最後に、距離(12)式に基づく共分散関数を構築する。本節 [1] の議論を踏まえ、ここでは、指数型共分散関数を適用することとする。距離(12)式に対して指数型共分散関数を適用するためには、距離(12)式がネットワーク上の最短経路距離になっていることを示せばよいが、本研究では(13)式の制約を導入していることから自明である。従って、(12)式で定義される距離もまた、拡張最短距離木を用いた本節 [1] の議論が適用できる。そのため、距離(12)式の共分散関数は、 dn と同様に、指数型共分散関数が適用できる。

[3] [1] で定義された dn に基づく共分散関数、及び、[2] で定義された dr 、 da に基づく共分散関数を結合することで、最短経路距離を考慮した任意地点間の共分散関数を定義する。ここでは、(14)式で定義される Separable 型共分散関数 (e.g. Cressie and Huang, 1999) を適用する。

$$C(d_1, d_2) = C(d_1)C(d_2) \quad (14)$$

d_1 及び d_2 は距離である。もし、共分散関数 $C(d_1)$ 及び $C(d_2)$ が共に共分散の十分条件を満足するならば、 $C(d_1, d_2)$ もまた十分条件を満足することが知られている (Cressie, 1993)。

Separable 型共分散関数の概念を用いて [1] 及び [2] で構築された共分散関数を結合することで、(15)式が与えられる。

$$C_{net}^*(d_{ij}) = \begin{cases} \tau^2 + \sigma^2 & (D_i = D_j \text{ and } s_i = s_j) \\ \sigma^2 \exp \left[- \left(\frac{|dr_{i,i} - dr_{j,i}|}{w_{dr}} + \frac{da_{i,i'}}{w_{da}} \right) \right] & (D_i = D_j \text{ and } s_i \neq s_j) \\ \sigma^2 \exp \left[- \left(\frac{dr_{i,j}}{w_{dr}} + \frac{dn_{i,j}}{w_{dn}} + \frac{dr_{j,j}}{w_{dr}} \right) \right] & (D_i \neq D_j) \end{cases} \quad (15)$$

$s.t. \ w_{dr} \leq (2/\pi)w_{da}$

(15)式を Kriging の基本式(1)式の共分散関数に仮定することで、鉄道ネットワークに起因した空間

的相互作用を空間従属性に考慮することができる。

4. 4. 最短経路距離及びユークリッド距離を考慮した共分散関数の提案

多くの既存研究で仮定されるように、ユークリッド距離に起因した空間従属性もまた考慮することが自然であろう (詳しくは 4 章 1 節)。そこで、前節で定義した最短経路距離に基づく共分散関数(15)式と通常のユークリッド距離に基づく共分散関数(4)式を結合することとする。類似の結合を行った空間統計学の先行研究として、Cressie et al. (2006) では、河川ネットワーク上の最短経路距離に基づく共分散関数式 C_{tree}^* 、及び、ユークリッド距離に基づく共分散関数 C_{euclid}^* を下式で結合し、これを用いた空間統計モデルを Kriging に適用することで、河川内の鉱物資源の量を予測している。

$$C^* = \lambda C_{tree}^* + (1 - \lambda) C_{euclid}^* \quad (16)$$

λ は両距離の空間従属性に対する説明力を表すパラメータであり、1 に近ければ最短経路距離が、0 に近ければユークリッド距離が、それぞれより強い説明力を持つことを意味する。なお、(16)式で定義される共分散関数が共分散の十分条件 (9-a), (9-b), (9-c') を満足することは簡単に示すことができる。

本研究では、(16)式を援用することで、交通ネットワークに基づく最短経路距離及びユークリッド距離の両方を考慮した共分散関数(17)式を用いる。

$$C^* = \lambda C_{net}^* + (1 - \lambda) C_{euclid}^* \quad (17)$$

$$s.t. \ w_{dr} \leq (2/\pi)w_{da}$$

ここで、 C_{net}^* は(15)式で示される共分散関数、 C_{euclid}^* は(4)式で示される共分散関数である。

5. 公示地価データを用いた実証

5. 1. 実証の概要

ユークリッド距離のみを考慮した従来の共分散関数(4)式を用いた Kriging [従来モデル]、及び、ユークリッド距離と最短経路距離の両方を考慮した共分散関数(17)式を用いた Kriging [提案モデル] をそれぞれ地価の予測に適用し、予測精度等の比較より、提案モデルの有用性を検証する。ここでは、提案モデル適用で必要となる駅勢圏は、ユークリッド距離

に基づくボロノイ分割で定義することとする。

本実証では、国土交通省の Web サイト、国土数値情報ダウンロードサービス (<http://nlftp.mlit.go.jp/ksj/index.html>) で提供されている『2007 年度住宅地公示地価データ』を使用する。対象範囲は東京都及び埼玉、千葉、神奈川の 3 県及び茨城県南部であり (図 4)、対象範囲内のデータ (地点) 数は 6,204 である。

被説明変数には公示地価 (万円 / m^2) の自然対数、説明変数には、最寄りの鉄道駅までの距離 [最寄駅距離]、最寄りの鉄道駅から東京駅までの鉄道ネットワーク上の最短経路距離 [東京距離]、所属する町丁目の人口密度 [人口密度] を用いることとする。

パラメータ推定には最尤法を用いる。また、予測精度の検証は、5-fold クロスバリデーション、即ち、サンプルをランダムに 5 つに分割し、5 分の 4 のサ

図 4 対象地域の公示地価の分布

ンプルを用いて残りの 5 分の 1 のサンプルの予測を行うことを、5 回繰り返すことによる予測精度の検証を行う。

以上の計算には、前述の統計ソフト CRAN (<http://cran.r-project.org/>) 提供のフリーの統計ソフトウェアである『R (バージョン 3.2)』を、64bit のメモリを持つ電子計算機上で操作することで行う。また、地図の描写、及び、最短経路距離の計算には ESRI 社 (<http://www.esri.com/>) 提供の『ArcGIS (バージョン 9.3)』を用いる。

5. 2. 実証の結果

5. 2. 1. パラメータの推定

各パラメータの推定結果を表 1 に示す。提案モデル及び従来モデルのトレンドパラメータの推定値は概ね類似した傾向を持つことがわかる。また、表内の w_{euclid} 、 w_{dn} 、 w_{da} 及び w_{dr} は各距離指標の増加に伴う空間従属性の及ぶ範囲の広さを表すパラメータであり、提案モデルの w_{dn} の推定値が他の w の推定値に比べ大きな値を示していることから、対象地域内における空間従属性が、鉄道ネットワークに沿って引き延ばされていることが確認できる。その一方で、 w_{da} と w_{dr} の推定値から、放射方向により強い空間従属性が認められる一方で、それらは (13) 式で示させる制約の影響を受けている、すなわち (13) 式の等号が成立していることが分かる。このことは、今回の地域に関しては、駅勢圏内の距離の定義が実際の現象に即していない可能性も考えられるため、例え

表 1 パラメータ推定結果

モデル	提案モデル		従来モデル	
パラメータ	推定値	t 値	推定値	t 値
定数項	12.97	595.1	12.97	489.9
最寄駅距離 (km)	-3.167×10^{-2}	96.80	-2.622×10^{-2}	65.97
東京距離 (km)	-1.119×10^{-1}	21.58	-1.135×10^{-1}	18.02
人口密度 (千人/ km^2)	1.163×10^{-5}	10.48	1.918×10^{-5}	14.23
τ^2	3.056×10^{-1}		3.836×10^{-1}	
σ^2	9.034×10^{-1}		9.736×10^{-1}	
w_{euclid}	39.39		39.42	
w_{dn}	51.97			
w_{dr}	2.047			
w_{da}	3.215			
λ	2.064×10^{-1}			
AIC	9524		10550	

ば、駅勢圏内の距離を単なるユークリッド距離で定義した場合との比較を通して、さらなる考察が必要である。 λ が約0.2であることは、空間従属性の2割が交通ネットワークで説明されることを示唆している。また、AICの比較より、提案モデルの予測能力が従来モデルを上回っていることが示唆される。

5.2.2. 予測精度の検証

平均二乗平方根誤差 ($RMSE$: Root Mean Square Error) を計算することで、実際の予測精度の比較を行う。 $RMSE$ は(18)式で求められる。

$$RMSE = \sqrt{\sum (y_0 - \hat{y}_0)^2 / 6204} \quad (18)$$

表2に $RMSE$ の計算結果を示す。表から、提案モデルの $RMSE$ は、従来モデルの $RMSE$ に比べて23%程度小さいことがわかる。このことから、交通ネットワークに起因した空間従属性を考慮することの重要性が示唆された。

提案モデルを用いた地価の予測結果を図5に、従来モデルを用いた同予測結果を図6にそれぞれプロットした。両予測結果を、鉄道ネットワークに沿った地域について比較すると、提案モデルの予測値は、従来モデルの予測値に比べて大きな数値を示す傾向があるとわかる。

次に、誤差率(19)式の差から、予測精度の差を地点別に比較する。

$$\text{誤差率} = \sqrt{\{(y_0 - \hat{y}_0) / y_0\}^2} \times 100 \quad (19)$$

両モデルの誤差率の差のプロットを図7に示す。この図から、提案モデルを用いることで予測精度が向上した地点の数は、精度の悪化した地点の数を明らかに上回っており、予測精度が向上した地点は、前述の鉄道ネットワークに沿って多く分布し、特に、東京中西部及び東京北東部の領域に多く分布していることが分かる。

表2 $RMSE$ 算出結果

	提案モデル	従来モデル
$RMSE$ (万円/ m^2)	6.392	8.333

図5 提案モデルの予測結果

図6 従来モデルの予測結果

図7 誤差率の差のプロット

6. 結論

本研究では、交通ネットワークに起因した空間的相互作用の結果として生じる空間従属性を明示的に考慮し、なおかつ空間統計学等の理論と整合的に任

意地点の空間内挿を行う新たな手法を、Kriging に用いられる共分散関数の拡張を行うことによって提案した。ここでは、都市/都市圏の巨視的な空間構造を決める交通ネットワークとして鉄道を考え、駅勢圏間の空間従属性を鉄道ネットワーク上の最短経路距離を用いた共分散関数によって表現した。一方、各駅勢圏内の微視的な構造が駅を中心とした放射環状道路ネットワークによって決まっていると仮定して、放射方向と環状方向の空間従属性を明示的に区別した共分散関数によって表現した。そして、二つの共分散関数を結合することで、都市/都市圏全域における交通ネットワークに起因する空間従属性を表現する共分散関数を導いた。さらにこれを、都市/都市圏全体における空間従属性のもう一つの要因である地点間のユークリッド距離に基づく従来の共分散関数と結合し、最短経路距離とユークリッド距離の両方に基づく共分散関数を提示した。

最後に、提案した共分散関数を用いた空間予測手法を東京都市圏の地価予測に適用し、従来手法による結果と比較することで、提案手法の予測能力の高さを確認し、空間的な予測において、交通ネットワークに起因する空間従属性を明示的に考慮することの重要性を確認した。

提案した手法では、必然的に予測値が駅勢圏の境界線上で不連続に変化することとなる点が一つの限界であるが、このことが予測結果にどのような影響を及ぼしているのかを考察するとともに、理論と整合する形でこれを避けることが可能か今後検討したい。一方で、本研究で提案した手法においてノード（駅）の数を無限にした極限を考えることにより、例えば河川ネットワークが存在する地理的空間における汚染物質の分布や、道路ネットワークが存在する地理的空間における車を起源とする騒音や汚染物質の分布の予測などへの応用が可能である。

また、Krivoruchko and Gribov (2004) の手法を用いた場合との予測結果の比較や、特に今回仮定した駅勢圏内における距離指標の妥当性の検討、さらには時間距離や費用距離等を用いた空間従属性の考慮について、引き続き実証的に試みたい。

謝辞

本稿の査読段階において、2名の匿名の査読者から、それぞれ非常に有益で示唆に富むコメントを頂いた。ここに記して感謝を申し上げる。無論、本稿に関する誤りは全て筆者らに属するものである。

参考文献

- 竹内伝史 (2000) 公共輸送の計画. 『交通工学 (土木教程選書)』, (青島縮次郎・本多義明・磯部友彦共著), 鹿島出版会, 185-222.
- 四茂野英彦. (1993) ネットワーク上での最短距離分布の計算可能性. 『GIS- 理論と応用』, **1**, 47-56.
- Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- Barry, R.D. and Ver Hoef, J.M. (1996) Blackbox kriging: spatial prediction without specifying the variogram. *Journal of Agricultural, Biological, and Environmental Statistics*, **1**, 297-322.
- Conley, T.G. and Topa, G. (2001) Socio-Economic Distance and Spatial Patterns in Unemployment. *Journal of Applied Econometrics*, **17**(4), 303-327.
- Cressie, N. (1993) *Statistics for Spatial Data*, Revised Edition, John Wiley & Sons.
- Cressie, N., Frey, J., Harch, B., and Smith, M. (2006) Spatial Prediction on a River Network. *Journal of Agricultural, Biological, and Environmental Statistics*, **11**(2), 127-150.
- Cressie, N. and Huang, H. C. (1999) Classes of Non-separable, Spatio-Temporal Stationary Covariance Functions. *Journal of the American Statistical Association*, **94**(448), 1330-1340.
- Curriero, F. C. (2006) On the Use of Non-Euclidean Distance Measures in Geostatistics. *Mathematical Geology*, **38**(8), 907-926.
- Dall'erba, S. (2005) Productivity convergence and spatial Dependence among Spanish regions. *Journal of Geographical Systems*, **7**(2), 207-227.
- Diggle, P. J. and Verbyla, A. P. (1998) Nonparametric Estimation of Covariance Structure in Longitudinal Data. *Biometrics*, **54**, 401-415.

- Garreta, V., Monestiez, P. and Ver Hoef, J. M. (2010) Spatial modeling and prediction on river networks: up model, down model or hybrid?. *Environmetrics*, **21**, 439-456.
- Gotway, C, A., Young, L, J., (2008) Geostatistics: What's Hot, What's Not, and Other Food for Thought. Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 8-16.
- Krivoruchko, K., Gribov, A. (2004) Geostatistical Interpolation and Simulation with Non-Euclidean Distances. *geoENV IV - Geostatistics for Environmental Applications (Quantitative Geology and Geostatistics)*, Kluwer Academic Publishers, Barcelona, Spain, 331-342.
- Tsutsumi, M. and Seya, H. (2009) Hedonic approaches based on spatial econometrics and spatial statistics: application to evaluation of project benefits. *Journal of Geographical Systems*, **11**(4), 357-380.
- Ver Hoef, J, M., Peterson, E., and Theobald, D. (2006) Spatial Statistical Models that Use Flow and Stream Distance. *Environmental and Ecological Statistics*, 449-464.
- Zimmerman, D. L. (1993) Another look at anisotropy in geostatistics. *Mathematical Geology*, **25**(4), 453-470.
- (2010 年 4 月 19 日原稿受理, 2011 年 5 月 19 日採用決定, 2011 年 6 月 16 日デジタルライブラリ掲載)