

空間計量経済学における近年の方法論的な発展

瀬谷 創[†]・泊 将史[†]

(受付 2024 年 7 月 3 日；改訂 8 月 21 日；採択 8 月 26 日)

要 旨

空間計量経済学という分野が産声を上げてから、もうすぐ50年が経過しようとしている。標準的な方法論は2000年代までにおおむね整備され、現在までに多くの実証研究が行われてきた。本稿では、空間計量経済学分野の近年の方法論的な発展に関する個人的な見解を論じる。具体的には、1) 識別と因果推論、2) 空間重み行列の特定化、3) 空間的自己相関構造の柔軟なモデル化、4) 時空間データのモデリング、5) ダーティデータのモデリングの観点から、近年の重要な方法論的な発展をレビューすることを試みる。

キーワード：空間計量経済学、空間重み行列、統計的因果推論、機械学習、ダーティデータ、時空間パネルデータ。

1. はじめに

1960～70年代、計量地理学の分野では、空間的自己相関(Spatial autocorrelation)が、最も基礎的かつ重要な問題のひとつと位置づけられ、空間データの分析モデリングのための知見が積み重ねられていった。その後、計量地理学の流れをくみながら、離散的な空間(市区町村等のゾーンなど)上の格子データ(Lattice data)のモデリングを扱う空間計量経済学(Spatial econometrics)と呼ばれる分野が、空間統計学と相互に影響を与えながらも異なる分野として発展した(Anselin, 2010; Arbia, 2011)。Anselin and Bera (1998)によると、「空間計量経済学」という用語は、1970年代初頭にベルギーの経済学者であるJean Paelinckが用い始めたものであり、Paelinck and Klaassen (1979)は空間計量経済学の分野における初のテキストとされている。その後、空間計量経済学の発展の一つの大きなきっかけになったのが、Anselin (1988)や、LeSage and Pace (2009) [各務・和合 訳, 2020]の出版であり、Anselin (1988)では、空間計量経済学が、「空間的自己相関と、空間的異質性(Spatial heterogeneity)を扱う計量経済学の一分野」と定義されている。

格子データのモデリングは、大きくBesag (1974)流のマルコフ確率場・条件付き分布に基づく方法と、Anselin (1988)流の非マルコフ確率場・同時分布に基づく方法に分類できる。モデルで言えば、前者はConditional autoregressive(CAR)モデルを用い、後者はSimultaneous autoregressiveモデルを用いることが多い。本稿は特に空間計量経済学に着目するため、説明は後者が中心となることをご容赦願いたい。なお、モデルの呼称については、統計コミュニティと空間計量経済学のコミュニティで異なる点があるので、注意が必要である(瀬谷・堤, 2014)。

空間計量経済学は2000年代に大きく花開いた分野である。その要因はいくつか考えられるが、私見では、1) 計量経済学者や統計学者の参入により、代表的なモデルの推定量の漸近的性

[†] 神戸大学 大学院工学研究科：〒657-8501 兵庫県神戸市灘区六甲台町1-1

質が整理され (Lee, 2023), モンテカルロ実験等によって小標本特性についても理解が進んだこと, 2)空間データの普及により, 空間的自己相関のモデル化のニーズが高まったこと, 3)計算ソフトの普及 (Anselin and Rey, 2022)や計算機性能の向上により手法実装の敷居が低下したこと, 4)地理情報システム (Geographic information system (GIS))の普及によって空間データの処理にかかるコストが大きく低下したこと, 5)Spatial Econometrics Associationの立ち上げ (Arbia, 2011)により, 集中的な議論ができる場が設けられたことなどが重要であったように思う. 当時, 数多くの国際誌で特集号が生まれ, Anselin (2010)は, それまでの30年間の回想論文において, 応用計量経済学や社会科学の方法論として, マージンからメインストリームへ移行したと述べている.

しかしその後, Gibbons and Overman (2012)が “Mostly pointless spatial econometrics?” という刺激的なタイトルの論文を地域科学分野のトップジャーナルの一つである Journal of Regional Science 誌の Symposium on spatial econometrics と題された特集号に発表した. これは, 因果の識別問題の観点から, 空間計量経済学の「モデルの関数形が既知であると仮定して推定値を計算し, モデル比較の手法を用いてモデルをデータに選択させるアプローチ」を批判したものである. 経済学では, 1990年代のいわゆる信頼性革命以降, 因果や識別が重要なトピックとなった. 一方空間計量経済学は, そのような動きに対して鈍感であったと言える. Mur (2013)は, 空間計量経済学が識別や因果関係に注意を払ってこなかったのは, おそらく初期の空間計量経済学の進展が, 時系列解析の進展の模倣によるものだったという歴史に関係していると指摘している. また, Pinkse and Slade (2010)は, “The future of spatial econometrics” と題するレビュー論文において, 1つの空間パラメータのみで空間的自己相関をモデル化する標準的な空間計量経済モデルについて, “laughable notion” と問題視している.

このように, 伝統的な空間計量経済学のアプローチについて, 特に2000年代後半頃から様々な課題が指摘されてきた. 一方で, それらに対処するためのアプローチの提案や, モデルの高度化も大きく進展した. しかし, 伝統的なアプローチを整理した堤・瀬谷 (2012a), 土木や地域科学関連分野の応用研究を整理した堤・瀬谷 (2012b)では, このような進展が反映されておらず, 近年の方法論的発展を和文で整理した文献は筆者らの知る限り存在しない. 以上の背景から本研究では, 1)識別と因果推論, 2)空間重み行列の特定化, 3)空間的自己相関構造の柔軟なモデル化, 4)時空間データのモデリング, 5)ダテデータデータのモデリングの観点から, 近年の重要な方法論的発展をレビューすることを試みる.

2. 空間計量経済モデルの概要

本研究では, 基本的にはクロスセクショナルな格子データにおける線形の空間計量経済モデルを対象とするが, 一部パネルデータも対象とする.

クロスセクションにおける代表的な空間計量経済モデルは,

$$(2.1) \quad \mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$(2.2) \quad \mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{u}, \mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\varepsilon}$$

$$(2.3) \quad \mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{W} \mathbf{X} \boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

$$(2.4) \quad \mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \mathbf{W} \mathbf{X} \boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

と定式化できる. ここで, サンプルサイズを N としたとき, \mathbf{y} は $N \times 1$ の被説明変数ベクトル, \mathbf{W} は $N \times N$ の対角要素を0とする空間重み行列, \mathbf{X} は $N \times K$ の説明変数行列, $\boldsymbol{\varepsilon}$ は $N \times 1$ の平均0, 均一分散の independent and identically distributed (i.i.d.) 誤差のベクトル (\mathbf{X} は定数項

を含むが、 $\mathbf{W}\mathbf{X}$ は定数項を含まないとする)、 β 、 γ は回帰係数に対応する $K \times 1$ 、 $(K-1) \times 1$ のパラメータベクトル、 λ と ρ は空間的自己相関の度合いを表すパラメータである。これらのモデルはそれぞれ空間ラグモデル (Spatial lag model) あるいは空間自己回帰モデル (Spatial autoregressive (SAR) model) (2.1)、空間誤差モデル (Spatial error model) (2.2)、SLX (Spatial lag of X) モデル (2.3)、空間ダービンモデル (Spatial Durbin model) (2.4) と呼称されることが多い。

3. 空間計量経済モデルの方法論的發展

空間計量経済学分野では、モデルの関数形と空間重み行列 \mathbf{W} が既知であると仮定して、ラグランジュ乗数検定や情報量規準などを用いてモデルをデータに選択させるアプローチが伝統的に採用されてきた (Anselin, 1988)。しかし、このようなアプローチに対して、識別の観点 (Gibbons and Overman, 2012) やモデル構造の柔軟性の観点 (Pinkse and Slade, 2010) から、批判が寄せられてきた (Debarsy and Le Gallo, 2024)。これらの概要について、それぞれ 3.1 節と 3.3 節で論じる。また、近年発展の著しい、空間重み行列の特定化 (3.2 節)、時空間データのモデリング (3.4 節)、ダティデータのモデリング (3.5 節) について、近年の方法論的な重要な発展をレビューする。

3.1 識別と因果推論

一定の条件を満たす \mathbf{W} (例えば, Kelejian and Piras, 2017) において、 $|\rho| < 1$ のとき、SAR モデルは、

$$(3.1) \quad \mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1} \mathbf{X}\beta + (\mathbf{I} - \rho\mathbf{W})^{-1} \boldsymbol{\varepsilon}$$

と誘導形で表現することができる。冪級数展開 (Power series expansion) により、 $(\mathbf{I} - \rho\mathbf{W})^{-1} = \mathbf{I} + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \dots$ が得られるため、 \mathbf{y} の条件付き期待値は、

$$(3.2) \quad E[\mathbf{y}|\mathbf{X}] = \mathbf{X}\beta + \rho\mathbf{W}\mathbf{X}\beta + \rho^2\mathbf{W}^2\mathbf{X}\beta + \dots$$

と表現できる。ここで、SEM は係数に $\gamma = -\lambda\beta$ という制約を置けば SDM に帰着する (瀬谷・堤, 2014)。したがって、第 2 章で導入したすべての空間計量経済モデルは、

$$(3.3) \quad E[\mathbf{y}|\mathbf{X}] = \mathbf{X}\beta + \mathbf{W}\mathbf{X}\pi_1 + \mathbf{W}^2\mathbf{X}\pi_2 + \dots$$

と誘導形で表現できることとなり ($\pi_r (r = 1, 2, \dots)$ はパラメータベクトル)、モデル間の差異は、 \mathbf{X} の何次の空間ラグ項を導入するかの違いに他ならなくなる。前述のとおり空間計量経済学では、LM 検定や情報量規準でモデル選択を行うアプローチをとってきた。ここでは \mathbf{W} が正しく特定化されているという前提が置かれている。しかし実際には真の \mathbf{W} は不明であり、 \mathbf{X} の高次の空間ラグ変数同士に (多くの場合比較的強い) 共線性がある中で、どのモデルがデータを発生させたかを識別することは困難である。これが Gibbons and Overman (2012) が批判した点の 1 つである。論文では、“these different specifications are generally impossible to distinguish without assuming prior knowledge about the true data generating process that we often do not possess in practice” と表現されている。

一方で、第 2 章で導入した空間計量経済モデル (式 (2.1)-(2.4)) は、それぞれ実証的なインプリケーションが大きく異なる。例えば、SAR モデルでは、 k 番目の説明変数 x_k の限界的な変化の影響が、他地域に $(\mathbf{I} - \rho\mathbf{W})^{-1}$ を通じて波及 (スピルオーバー) するが、SEM では通常の線形回帰モデル同様にこのような効果はない。SLX モデルは、1 次の空間ラグ項のみが導入されているため、波及効果はローカルである。したがって、本来はモデル特定化の背景として

理論的な考察が必要であるが、空間計量経済モデルを用いた実証研究では、その部分が捨象され、データによる統計的なモデル選択のみに頼ったモデル特定化が行われている例が少なくない (SAR モデルのゲーム理論的な解釈については Lee, 2023 を、社会的相互作用的な解釈については石井 他, 2018) を参照されたい). Pointless という誹りを免れるためには、モデル特定化のための考察を厳密に行う必要があり、少なくとも目的に応じた使い分けを行う必要があるといえよう。

まず、目的が除外変数バイアスの緩和にある場合、誤差相関を考慮する SEM を用いることが自然である。実証研究では、空間的な固定効果 (例えば地価のモデリングにおける市区町村ダミー) の導入により空間的な相関を持つ除外変数に対処することが多いが (Kuminoff et al., 2010), どのような空間単位を選定するか (例えば、町丁目 or 市区町村) に結果が依存するという課題がある。空間単位が粗すぎるとコントロールとして不十分であるが、詳細に取りすぎると興味のある効果が抽出できなくなってしまう (Abbott and Klaiber, 2011; von Graevenitz and Panduro, 2015)。したがって、SEM によるコントロールは (W の特定化の問題は残るが) 実証研究において使いやすい (Anselin and Arribas-Bel, 2013)。他のアプローチとして、von Graevenitz and Panduro (2015) は、Generalized additive model (GAM) のフレームで緯度経度の関数を導入し、感度分析的な考察を行うことを提案している。興味のある変数が地理的な境界で非連続な場合は (例えば税率が異なるなど)、境界でモデルの空間的な差分を取ることで除外変数を difference-out する Spatial differencing と呼ばれる方法も使用可能である (例えば、Holmes, 1998; Belotti et al., 2018; Klein and Tchuente, 2023)。

次に、隣接地域における説明変数からのスピルオーバー効果を分析したい場合を考える。空間計量経済学の普及に大きな影響を果した LeSage and Pace (2009) が推奨したこともあり、現在までに SDM を用いた数多くの実証研究が行われてきた (例えば、Seya et al., 2012)。しかし、Gibbons and Overman (2012) は、前述した識別問題のため、SDM や SAR ではなく、誘導系の SLX を用いることを推奨している。Halleck Vega and Elhorst (2015) は、SAR と異なる SLX の利点として、 W をパラメタライズして、局所最適に陥らずに最尤法によりパラメータを推定できる点を挙げている (Olmo and Sanso-Navarro, 2023 はさらに柔軟なセミパラメトリックな SLX モデルを提案している)。ただし、 WX が外生性を満たすか否かについても、別途慎重な検討が必要である。例えば、Lyytikäinen (2012) は、政策の変更を外生的変動として、 WX の操作変数に用いている。

内生効果 ρ 自体に興味がある場合、SDM や SAR モデルを推定する必要がある。ここで Wy は内生変数であるため、(疑似)最尤法 (Lee, 2004) や操作変数法 (Kelejian and Prucha, 1998) を用いてパラメータ推定を行うが、 X や WX の内生性を考慮する場合には後者が用いられる。 Wy に対する操作変数 (instrumental variable (IV)) としては、 X の高次の空間ラグ変数 [WX, W^2X, W^3X, \dots] が用いられることが多い。このようなアプローチについて、Gibbons and Overman (2012) は、1) X の空間ラグ変数が除外制約を満たさない場合、IV として適切でないこと、2) X の空間ラグ変数は共線性の問題から、独立した外生的変動をわずかしき提供できず (特に SDM の場合)、いわゆる weak instruments の問題が存在することを指摘した。実際に空間計量経済モデルの実証研究では、 X の空間ラグ変数が除外制約を満たすか否かに関する考察が行われることは稀である。 X の空間ラグ変数が IV として適切でない場合、別の IV を準備する必要がある。例えば、政策の変更を外生的変動として利用した前述の Lyytikäinen (2012) や、制度上の閾値を利用した Fruehwirth (2013) 等の取り組みが参考になると思われる。

さて、経済学では、1990 年代のいわゆる「信頼性革命」以降、統計的な意味での因果推論が重要になった。因果推論の方法論は、大きく Pearl 流の Directed acyclic graph (DAG) アプローチ (Pearl and Mackenzie, 2018 [夏目 訳, 2022]) と、Rubin 流の潜在的結果 (Potential outcome) ア

プローチ (Imbens and Rubin, 2015) に大別できる。経済学では、特に後者を実証分析に用いることが一般化した。本稿では、後者のアプローチにおける空間計量経済モデルの可能性について論じる (前者と空間モデルの関係については、Akbari et al., 2023 等を参照されたい)。

潜在的結果アプローチで平均処置効果を求める上では、Stable Unit Treatment Value Assumption (SUTVA) 条件が満たされる必要がある (Rubin, 1980)。SUTVA 条件は、1) No interference between units (潜在的結果が他の主体 (地域) の処置状態に依存しない) と 2) No multiple versions of treatments (ある処置に異なる形 (バージョン) は存在しない) からなる。ここで仮定 1 は、スピルオーバー効果や一般均衡的效果を除外する条件であるが、現実の処置においてはこれが成り立たない場合も少なくない。例えば、都市の文脈では社会資本が隣接自治体に与える正の影響や (要藤・吉村, 2016)、都市計画の用途地域の変更が近隣地域の住宅建設数に与える負の影響などが挙げられる (Greenaway-McGrevy and Phillips, 2023)。相互干渉 (Interference) が存在する場合、ある主体は実際には処置を受けていないにも関わらず処置効果を間接的に受けることになるため、これを無視して通常の Difference-in-differences (DID) 推定を行うと、平均処置効果の推定値にバイアスが生じる。

相互干渉をモデルで扱う標準的なアプローチは、処置のスピルオーバー効果を低次元の十分統計量に変換する Exposure mapping と呼ばれる方法であり (Aronow and Samii, 2017)、処置された隣接主体のシェア (Hudgens and Halloran, 2008) などがよく用いられる。また、空間的なスピルオーバーの場合、隣接地域、隣々接地域における処理の有無 (要藤・吉村, 2016) や隣接地域へのダミー変数の導入 (Butts, 2023) などが用いられている。Exposure mapping の方法の一つとして、空間重み行列を用いることができる (Delgado and Florax, 2015; Bardaka et al., 2019)。今、 \mathbf{Y} を $NT \times 1$ の被説明変数のベクトル (N は地域数、 T は観測時点数)、 \mathbf{D} を $NT \times 1$ の処置群ベクトル (処置群であれば 1、対照群であれば 0)、 \mathbf{T} を $NT \times 1$ の処置後ベクトル (処置後であれば 1、処置前であれば 0)、 \mathbf{I}_{NT} を NT 次元の単位行列、 ι を 1 からなる $NT \times 1$ のベクトル、 ε を $NT \times 1$ の平均 0、均一分散の i.i.d. 誤差のベクトルとする。空間重み行列が時間不変であるとき、 NT 次元の空間重み行列 \mathbf{W}_{NT} は、 $\mathbf{W}_{NT} = \mathbf{I}_T \otimes \mathbf{W}_N$ と与えることができる (\otimes はクロネッカー積)。このとき、SLX 型の Spatial DID モデル (Delgado and Florax, 2015) は、

$$(3.4) \quad \mathbf{Y} = \alpha_0 \iota + \alpha_1 \mathbf{D} + \alpha_2 \mathbf{T} + \alpha_3 (\mathbf{I}_{NT} + \phi \mathbf{W}_{NT}) \mathbf{D} \circ \mathbf{T} + \varepsilon$$

と定式化できる。ここで、 $\alpha_k (k = 1, 2, 3)$ はパラメータ、 ϕ は空間パラメータ、 \circ はアダマール積である。 $\alpha_4 = \alpha_3 \phi$ と置き直したとき、 α_3 は直接効果 (Average treatment effects on the treated (ATT)) を、 α_4 は間接効果 (スピルオーバー効果) を表す (Hudgens and Halloran, 2008)。Bardaka et al. (2019) は、Delgado and Florax (2015) を複数種類の処置がある場合に、Chagas et al. (2016) は複数の空間重み行列がある場合に、それぞれ拡張している。Hudgens and Halloran (2008) における処置された隣接主体のシェアは、 \mathbf{W}_N が行規準化された場合に相当する。また、Chagas et al. (2016) のように複数の重みを用いれば、要藤・吉村 (2016) のように隣接地域、隣々接地域を考慮することもできる。

因果推論における空間モデルの活用に関するさらなる話題については、Kolak and Anselin (2020)、Reich et al. (2021)、Debarsy and Le Gallo (2024) などを参照されたい。

3.2 空間重み行列の特定化

前節では、識別の観点から、SLX モデルの有用性について議論した。しかし、どのモデルを用いるにせよ、基本的には空間重み行列 \mathbf{W} の特定化が必要である。Stakhovych and Bijmolt (2009) は、 \mathbf{W} の与え方を、(1) 完全に外生とする、(2) データから決定する、(3) 推定するとい

う3つに分類した。(1)は、地域の境界が接しているか否か(隣接行列)や、距離の逆数、ドロネー三角網等で与える典型的な方法である。(2)には、情報量規準 (Seya et al., 2013; Agiakloglou and Tsimpanos, 2023)やベイズモデル平均化 (Debarsy and LeSage, 2022)等のアプローチが存在する(\mathbf{W} の特定化については、Griffith, 2020等を参照されたい)。

ここで、(1)のアプローチがよく用いられる理由としては、「距離が近い事物はより強く影響しあう」という地理学の第一法則(first law of geography)が持ち出されることが多いが(Anselin, 1988), \mathbf{W} の外生性が自然に担保されるという計量経済学的な側面も重要である。すなわち、SARモデルにおける k 番目の説明変数の限界効果は、 $\partial \mathbf{y} / \partial \mathbf{x}'_k = \beta_k (\mathbf{I} - \rho \mathbf{W})^{-1}$ で与えられるため、処置が \mathbf{W} を変化させるような内生性が存在する場合、限界効果にバイアスが発生する。しかし、経済理論的な理由から \mathbf{W} が社会経済的な距離として与えられる場合もあり(例えば、Behrens et al., 2012), \mathbf{W} の内生性を考慮した推定法が必要である。

Qu and Lee (2015)は、重み行列の内生性を考慮する実用的な方法を提案した。まず、 p 個の内生変数 $\mathbf{z} = (z_1, \dots, z_p)'$ を用いて、距離測度を $f(\cdot)$ としたとき、重み行列を $w_{ij} = f(z_i, z_j)$ と表現する(i, j は地域を表す添字)。例えば、 $p = 2$ で z_1 がGDP、 z_2 が人口だとすると、重み行列はGDPおよび人口の類似度を意味する。ここで、 \mathbf{z} が外生説明変数と誤差項の線形モデルで与えられると仮定すると、

$$(3.5) \quad \mathbf{Z} = \mathbf{X}_2 \mathbf{B} + \mathbf{E}$$

とおける。ここで \mathbf{Z} は $N \times p$ の内生変数の行列、 \mathbf{X}_2 は $N \times q$ の外生説明変数からなる行列、 \mathbf{B} は $q \times p$ の回帰係数行列、 \mathbf{E} は $N \times p$ の誤差項の行列である。Qu and Lee (2015)は、SAR方程式(式(2.1))の誤差項 ε と、 \mathbf{W} の要素に関する方程式の誤差項 \mathbf{E} が相関を持つことを許容し、QML等による複数のパラメータ推定方法を提案した。結果として、SAR方程式は、残差 $\mathbf{Z} - \mathbf{X}_2 \mathbf{B}$ をコントロール変数として導入する形で再構成される。また、 ε と \mathbf{E} の相関を調べれば、 \mathbf{W} の内生性についての仮説検定も可能である。本アプローチは、 w_{ij} がバイラテラル(bilateral)変数(例えば地域間流動など)で与えられる場合や(Qu et al., 2021)、時間変化する \mathbf{W}_t の場合にも拡張されている(Qu et al., 2017)。

瀬谷・堤(2014)の時点では、「 \mathbf{W} の要素の推定に関する研究は、まだまだ発展途上であるため、現状では(1)や(2)のアプローチを用いて作成された複数の \mathbf{W} から、何らかの基準により、最適なものを選択するというアプローチが重要になろう。」と指摘した。しかし近年、 \mathbf{W} の推定に関する研究も蓄積されつつある。ただし、サンプルサイズ N に対して \mathbf{W} の要素数は(対角項を除けば) $N^2 - N$ 存在するため、多くの場合推定には $T \gg N$ が要求され、 T が小さい場合への適用性をモンテカルロ実験で調べる形がとられている。既往研究のアプローチは大きく、1)機械学習の方法と2)ベイズ推定を用いた方法に区分できる。前者は、モデルの適合度と \mathbf{W} の疎性のバランスを調整する方法であり、Lam and Souza (2020)のAdaptive LASSO, de Paula et al. (2024)のAdaptive elastic net generalized methods of moments(GMM) (Caner and Zhang, 2014)を用いた研究が挙げられる。交通分野の交通状態推定も、LASSOによる \mathbf{W} の推定と類似するところがある(原他, 2016; Hara et al., 2018)。一方後者は、 \mathbf{W} に事前分布を与えるアプローチである。Krisztin and Piribauer (2023), Piribauer et al. (2023)は、 \mathbf{W} の非対角要素に事前分布としてベルヌーイ分布を仮定し、また近傍となる主体の数がベータ二項分布に従うと仮定することで、ハイパーパラメータの設定によって \mathbf{W} の疎性がコントロールできることを示した。この手法は、空間データ分析にとって標準的な、 $N > T$ の状況での適用も可能である。事前分布を用いた方法には、他にもCARモデルをベースとしたGao and Bradley (2019)等が存在する。このように機械学習やベイズ推定を応用したいくつかの手法が提案されてきているが、 \mathbf{W} の推定方法は未だ確立されてはいない。

3.3 空間的自己相関構造の柔軟なモデル化

典型的な空間計量経済モデルのシンプルなモデル構造については、解釈の容易さ等の利点はあるが、批判も少なくない (Pinkse and Slade, 2010). しかし近年、非線形性や異質性などを考慮した、より柔軟なモデルが、機械学習アプローチを取り入れながら提案されている. 本節では、それらの研究を概観する.

このような方向での萌芽的な取り組みとして、Pinkse et al. (2002)が挙げられる. Pinkse et al. (2002)は、SAR タイプのモデル

$$(3.6) \quad \mathbf{y} = \mathbf{W}(g)\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

において、 \mathbf{W} の要素を $w_{ij} = g(d_{ij})$ と与えた. $g(\cdot)$ は、距離 d_{ij} の影響度合いを表す未知の関数であり、例えばフーリエ基底関数や多項式基底関数で近似すればよい. Sun (2016)は、Pinkse et al. (2002)のアプローチを、 $w_{ij} = g(z_{ij})$ と内生変数に依存する場合に拡張した. ただし、説明変数が複数の場合には、前述の Qu and Lee (2015) アプローチ等を用いる必要がある. 若干異なるアプローチとして、塚井・奥村 (2007), Sun (2024)は、 $\rho(z)\mathbf{W}\mathbf{y}$ と functional coefficient 型の定式化を行った. 前者は z が外生変数と仮定している一方、後者では z の内生性が考慮されている. Hoshino (2022)は、 $\mathbf{W}g(\mathbf{y})$ と非線形の内生効果を考慮し、sieve IV によるパラメータ推定法を提案している. Mínguez et al. (2022)は、SAR タイプの空間パネルモデルにおいて、説明変数の非線形性を考慮しつつ、空間座標の非線形効果を導入するモデルを考案している. Geniaux and Martinetti (2018)は、Geographically weighted regression (GWR) モデルと SAR モデルを結合し、空間パラメータを場所によって変えるモデルを構築した(ただし、GWR モデルはローカルな多重共線性に弱いため注意が必要である、村上, 2024).

さらに柔軟な関数系として、グラフニューラルネットワークを用いた研究も行われている. Zhu et al. (2022)は、活性化関数を σ としたとき、 $\mathbf{y} = \sigma(\tilde{\mathbf{W}}\mathbf{X}\Theta)$ とするモデル化を行っている (Θ はニューラルネットワークにおける重み、 $\tilde{\mathbf{W}}$ は \mathbf{W} から構築されたグラフラプリアン行列である. 簡単のため中間層とバイアス項は無視している). Yang et al. (2022)は、不動産価格予測の文脈で、予測地点を含む形で $\tilde{\mathbf{W}}$ と \mathbf{X} を構築し、予測地点においては \mathbf{y} が欠損しているとするモデル化のほうが、学習地点のみでモデルを構築し ($\hat{\Theta}$ を得て)、予測地点のみで構築された $\tilde{\mathbf{W}}$ と \mathbf{X} に適用するよりも予測正確度が高いと指摘している. 類似の取り組みには、Wang and Song (2023)と Xiao et al. (2023)があり、それぞれモデル化の方法が若干異なる.

3.4 時空間データのモデリング

Elhorst (2022)は、次式に示す一般化された空間パネル(データ)モデルを導入し、各項の役割について説明を行っている.

$$(3.7) \quad \begin{aligned} \mathbf{y}_t &= \tau\mathbf{y}_{t-1} + \rho\mathbf{W}\mathbf{y}_t + \eta\mathbf{W}\mathbf{y}_{t-1} + \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{X}_t\boldsymbol{\gamma} + \Sigma_r\boldsymbol{\Gamma}_r'f_{rt} + \mathbf{u}_t \\ \mathbf{u}_t &= \lambda\mathbf{W}\mathbf{u}_t + \boldsymbol{\varepsilon}_t \end{aligned}$$

ここで、 \mathbf{y}_t は被説明変数 y_{it} からなる $N \times 1$ ベクトル、 \mathbf{y}_{t-1} 、 $\mathbf{W}\mathbf{y}_t$ 、 $\mathbf{W}\mathbf{y}_{t-1}$ はそれぞれ \mathbf{y}_t の時間、空間、時空間ラグ変数、 \mathbf{X}_t は $N \times K$ の説明変数行列、 $\mathbf{W}\mathbf{X}_t$ はその空間ラグ変数である. また、 τ 、 ρ 、 η 、 λ はそれぞれの変数の感応度を示すパラメータであり、 $\boldsymbol{\beta}$ 、 $\boldsymbol{\gamma}$ はそれぞれ $K \times 1$ 、 $(K-1) \times 1$ の回帰係数ベクトルである. モデルに $\mathbf{W}\mathbf{X}_{t-1}$ が導入されていない理由は、この項が $\mathbf{W}\mathbf{y}_{t-1}$ に含まれているため、識別の問題を引き起こすためである (Anselin et al., 2008, Anselin 2021). 式(3.7)のモデルは、時間方向のラグが入っているという意味で、動学的空間パネルモデルとなっている. 瀬谷・堤 (2014)では、いくつかの静学的空間パネルモデルが説明されているが、2000年代後半以降、動学的な空間パネルモデルの推定に関する研究が大き

く発展した(詳しくは, Lee, 2023 を参照されたい).

もう一つの近年の重要な進展が, $\Sigma_r \Gamma_r' f_{rt}$ 項の存在である. 空間計量経済モデルでモデル化される空間的自己相関は, \mathbf{W} を通したローカルな相関である. 一方で, 例えば東京と大阪の類似性など, 地理的な距離に依存しない形でのグローバルなクロスセクショナル相関の存在も考えられる (Pesaran, 2015). $\Sigma_r \Gamma_r' f_{rt}$ 項は, このようなグローバルなクロスセクショナル相関を考慮するために導入されるものである. Elhorst (2024)によれば, ローカル・グローバルの両方のクロスセクショナル相関を導入したモデルはまだまだ少ない. ベンチマーク的な研究である Shi and Lee (2017)は, 式(3.7)から $\mathbf{W} \mathbf{X}_{t\gamma}$ を除いたモデルの QML 推定量を導出し, その一致性を証明している. なお, 通常の空間パネルモデルは N が大きく, T が小さいという条件を仮定したものが多く, $\Sigma_r \Gamma_r' f_{rt}$ 項を導入したモデルは, 通常 T も大きいことが仮定されていることに注意されたい.

$\Sigma_r \Gamma_r' f_{rt}$ における $f_{rt}(r = 1, \dots, R)$ は, Bai (2009)の相互作用固定効果(interactive fixed effects)モデルで用いられたものであり, 因子分析における共通因子を意味する(和文の解説としては, 奥井, 2014 が参考になる). 例えば, 共通因子が $f_{1t} = (1, \dots, 1)'$, $f_{2t} = (\xi_1, \dots, \xi_T)'$ の2つであるとき, それぞれの負荷量を $\Gamma_1 = (\delta_1, \dots, \delta_N)'$, $\Gamma_2 = (1, \dots, 1)'$ で与えれば, $\Sigma_r \Gamma_r' f_{rt}$ 項は標準的な時間と主体(空間)の固定効果に等しくなる. 無論, このように時間効果と空間効果を分離する必要はなく, 相互作用固定効果を用いれば, 因子分析により場所によって異なる時間トレンドや, クロスセクショナルな相関を捉えることができる. また, $\Sigma_r \Gamma_r' f_{rt}$ 項を, (空間固定効果を導入しつつ), 被説明変数やその時間ラグ, さらには必要に応じて説明変数のクロスセクション平均で置き換える簡易推定法も発展している (Pesaran, 2015; 千木良他, 2011). このように, 空間パネルデータ分析のモデルは近年, より一般化される方向で発展したが, Angrist and Pischke (2009) [大森 他 訳, 2013]で議論されている通り, 実証分析において時間ラグを入れるかどうかの選択は難しい問題であり, 異なる識別の仮定を用いた感度分析的な考察による頑健性のチェックが極めて重要である.

その他の近年開発された空間計量経済学に関連する時空間モデルを以下で概説する. Amba and Le Gallo (2022)は, 空間パラメータが時間変化することを許容した空間パネルモデルを構築している. 一方, LeSage and Chih (2018), Aquaro et al. (2021)は, 回帰係数と空間パラメータが場所によって異なる Heterogeneous coefficients 型のモデルを構築した(ただし, T が大きいことが必要). Chen et al. (2022)は, そこに共通因子を導入し, Bao and Zhou (2023)は動学的パネルに拡張している. Wu and Matsuda (2021)は, パラメータを閾値変数で分類し, 複数のレジームを考慮する動学的空間パネルモデルを構築した. Yang and Lee (2019)は, 同次方程式タイプの動学的空間パネルモデルを構築し, QML 推定量と IV 推定量の漸近的性質について議論している.

Otto et al. (2023, 2024a), は, それぞれ空間パネルタイプの確率的ボラティリティモデルと Autoregressive conditional heteroscedasticity (ARCH)モデルを構築している (Otto et al., 2024b のレビュー参照). 塚井・小林 (2007)は, 社会資本の長期記憶性を考慮するために, SLX タイプの Auto-regressive fractionally integrated moving averaged model with exogenous variables (ARFIMAX)モデルを開発した. Cho et al. (2023)は, 時空間の自己回帰分布ラグ (Autoregressive distributed lag (ARDL))モデルについてのレビューを行っている. Ermagun and Levinson (2018)は, 交通量の時空間予測の文脈で, Space-time autoregressive integrated moving average (STARIMA)モデルおよびその関連モデルの可能性を丹念にレビューしている. Elhorst et al. (2021)は, Vector autoregressive (VAR)モデルと空間計量経済モデルの関係性について整理している.

3.5 ダーティデータのモデリング

空間計量経済学で研究されているほとんどのモデルは(A)データ欠損がなく、(B)観測位置が正確というクリーンな空間データを仮定している。しかし、利用可能な空間データではこれらが満たされないことが多く、ダーティデータに対する空間モデルの開発が求められている(Arbia et al., 2016)。なお、本節のレビューは離散空間における格子データだけでなく、連続空間における地球統計データも対象に含んでいる。

まず、(A)のデータ欠損への対応策として、最も標準的に用いられている手法は、欠損を含むデータは捨て去るという、リストワイズ除去と呼ばれる手法である。King et al. (2001)によれば、93~97年に調査した政策科学系の雑誌では、94%の研究で同手法による対応が行われていた。しかし、リストワイズ除去が有効なのは、Missing completely at randomが成り立つ場合、すなわち欠損するかどうかは被説明変数の値や説明変数の値に依存せず、完全にランダムであるという、限定的なケースのみである。特に空間計量経済モデルの場合、データの除去は空間重み行列を変化させるため、注意が必要である(Tsutsumi and Seya, 2009)。LeSage and Pace (2004)およびWang and Lee (2013)は、欠損が説明変数に依存するMissing at randomな状況において、それぞれExpectation maximization (EM)タイプのアルゴリズムと非線形最小二乗法を用いたパラメータ推定方法を提案している。欠損が被説明変数の値に依存するMissing not at random(あるいはNon-ignorable missing)のケースに対しては、サンプルセレクションを考慮した空間計量経済モデル(社会的相互作用モデル)が提案されている(Hoshino, 2019; Seya et al., 2021; Bao et al., 2024)。転移学習によるバイアス補正の検討も進みつつある(Zeng et al., 2024)。また、異なるアプローチとして、Arbia et al. (2022)はギブズサンプラーによる補正を、Arbia and Nardelli (2024)は、フォーマルなサンプリングを模倣するような再サンプリングを提案している。なお、Missing not at randomの状況で、例えば上側の分位点を中心にサンプリングがなされている場合(例えば、インフラの維持管理において危険性の高い箇所を中心に観測が行われているなど)、空間パターンにも歪みが生じ、空間内挿結果にGranularity biasと呼ばれるバイアスが生じる(Shirota and Gelfand, 2022)。これについては、空間計量経済モデルにおいても同様の注意が必要となる(Baldoni et al., 2023)。

次に、(B)の位置誤差については、まずGlobal navigation satellite system (GNSS)のようにそもそも位置情報に観測誤差が含まれる場合が考えられる。また、個人情報保護のための位置情報のマスキング(ジオマスキング)により位置誤差が生じるケースもある。Santi et al. (2021b)は、空間計量経済モデルとともにジオマスキングをマーク付き点過程としてモデル化し、ジオマスキングの影響を考慮した尤度最大化を行うことを提案した。Santi et al. (2021a)は、さらに位置情報が存在しない場合でも残差の空間的自己相関の影響を緩和するために、属性の類似性を用いて空間重み行列を構築することを提案した。

4. おわりに

Debarys and Le Gallo (2024)は、経済誌の上位5誌に掲載された論文のうち、「空間計量経済学」モデルを利用したものは、今やほとんどネットワーク計量経済学の文献にしかない指摘している。その一因として、Gibbons and Overman (2012)が指摘したように、空間計量経済学が識別や因果関係に十分な注意を払ってこなかったことや、Pinkse and Slade (2010)が述べた通り、空間計量経済学の標準的なモデルが、“laughable notion”と指摘されるほどにシンプルであったことが考えられる。

一方で、本研究でレビューした通り、前者・後者両方について、2000年代後半頃から様々な進展が見られた。本研究における整理が、上述したような状況の解消に少しでも役に立てば幸

いである。ただし本稿では、CARタイプのモデルや、一般化線形モデル、非線形モデル、さらには実装面(ソフトウェア等)については触れることができなかった。これらについては、別の機会に整理を試みたい。

謝 辞

本研究は JSPS 科研費 24K00175, 24K00997, 24K01004 の助成を受けたものです。

参 考 文 献

- Abbott, J. K. and Klaiber, H. A. (2011). An embarrassment of riches: Confronting omitted variable bias and multi-scale capitalization in hedonic price models, *Review of Economics and Statistics*, **93**(4), 1331–1342.
- Agiakloglou, C. and Tsimpanos, A. (2023). Evaluating the performance of AIC and BIC for selecting spatial econometric models, *Journal of Spatial Econometrics*, **4**, 2, <https://doi.org/10.1007/s43071-022-00030-x>.
- Akbari, K., Winter, S. and Tomko, M. (2023). Spatial causality: A systematic review on spatial causal inference, *Geographical Analysis*, **55**(1), 56–89.
- Amba, M. C. and Le Gallo, J. (2022). Specification and estimation of a periodic spatial panel autoregressive model, *Journal of Spatial Econometrics*, **3**, 13, <https://doi.org/10.1007/s43071-022-00028-5>.
- Angrist, J. D. and Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton. (大森義明, 小原美紀, 田中隆一, 野口晴子 訳 (2013). 『「ほとんど無害」な計量経済学—応用経済学のための実証分析ガイド』, NTT 出版, 東京.)
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- Anselin, L. (2010). Thirty years of spatial econometrics, *Papers in Regional Science*, **89**(1), 3–25.
- Anselin, L. (2021). Spatial models in econometric research, *Oxford Research Encyclopedia of Economics and Finance*, Oxford University Press, Oxford, <https://doi.org/10.1093/acrefore/9780190625979.013.643>.
- Anselin, L. and Arribas-Bel, D. (2013). Spatial fixed effects and spatial dependence in a single cross-section, *Papers in Regional Science*, **92**(1), 3–18.
- Anselin, L. and Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics, *Handbook of Applied Economic Statistics* (eds. A. Ullah and D. E. A. Giles), Marcel Dekker, New York.
- Anselin, L. and Rey, S. J. (2022). Open source software for spatial data science, *Geographical Analysis*, **54**(3), 429–438.
- Anselin, L., Le Gallo, J. and Jayet, H. (2008). Spatial panel econometrics, *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice* (eds. B. H. Baltagi, Y. Hong, G. Koop, W. Krämer and L. Matyas), Springer, Berlin, Heidelberg.
- Aquaro, M., Bailey, N. and Pesaran, M. H. (2021). Estimation and inference for spatial models with heterogeneous coefficients: An application to US house prices, *Journal of Applied Econometrics*, **36**(1), 18–44.
- Arbia, G. (2011). A lustrum of SEA: Recent research trends following the creation of the spatial econometrics association (2007–2011), *Spatial Economic Analysis*, **6**(4), 377–395.
- Arbia, G. and Nardelli, V. (2024). Using web-data to estimate spatial regression models, *International*

- Regional Science Review*, **47**(2), 204–226.
- Arbia, G., Espa, G. and Giuliani, D. (2016). Dirty spatial econometrics, *The Annals of Regional Science*, **56**, 177–189.
- Arbia, G., Matsuda, Y. and Wu, J. (2022). Estimating spatial regression models with sample data-points: A Gibbs sampler solution, *Spatial Statistics*, **47**, <https://doi.org/10.1016/j.spasta.2021.100568>.
- Aronow, P. M. and Samii, C. (2017). Estimating average causal effects under general interference, with application to a social network experiment, *The Annals of Applied Statistics*, **11**(4), 1912–1947.
- Bai, J. (2009). Panel data models with interactive fixed effects, *Econometrica*, **77**(4), 1229–1279.
- Baldoni, E., Coderoni, S. and Esposti, R. (2023). The productivity-environment nexus in space. Granularity bias, aggregation issues and spatial dependence within Italian farm-level data, *Journal of Cleaner Production*, **415**, <https://doi.org/10.1016/j.jclepro.2023.137847>.
- Bao, Y. and Zhou, X. (2023). Heterogeneous spatial dynamic panels with an application to US housing data, *Spatial Economic Analysis*, **18**(2), 259–285.
- Bao, Y., Li, G. and Liu, X. (2024). A spatial sample selection model, *Oxford Bulletin of Economics and Statistics*, **86**(4), <https://doi.org/10.1111/obes.12599>.
- Bardaka, E., Delgado, M. S. and Florax, R. J. (2019). A spatial multiple treatment/multiple outcome difference-in-differences model with an application to urban rail infrastructure and gentrification, *Transportation Research Part A: Policy and Practice*, **121**, 325–345.
- Behrens, K., Ertur, C. and Koch, W. (2012). ‘Dual’ gravity: Using spatial econometrics to control for multilateral resistance, *Journal of Applied Econometrics*, **27**(5), 773–794.
- Belotti, F., Di Porto, E. and Santoni, G. (2018). Spatial differencing: Estimation and inference, *CESifo Economic Studies*, **64**(2), 241–254.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion), *The Journal of the Royal Statistical Society, Series B*, **36**, 192–236.
- Butts, K. (2023). Difference-in-differences estimation with spatial spillovers, arXiv preprint, <https://doi.org/10.48550/arXiv.2105.03737>.
- Caner, M. and Zhang, H. H. (2014). Adaptive elastic net for generalized methods of moments, *Journal of Business & Economic Statistics*, **32**(1), 30–47.
- Chagas, A. L., Azzoni, C. R. and Almeida, A. N. (2016). A spatial difference-in-differences analysis of the impact of sugarcane production on respiratory diseases, *Regional Science and Urban Economics*, **59**, 24–36.
- Chen, J., Shin, Y. and Zheng, C. (2022). Estimation and inference in heterogeneous spatial panels with a multifactor error structure, *Journal of Econometrics*, **229**(1), 55–79.
- 力石真, 瀬谷創, 福田大輔 (2018). 社会的相互作用に着目したエビデンスベース研究の展開と土木計画への応用可能性, *土木学会論文集 D3(土木計画学)*, **74**(5), I_715–I_734.
- 千木良弘朗, 早川和彦, 山本拓 (2011). 『動学的パネルデータ分析』, 知泉書館, 東京.
- Cho, J. S., Greenwood-Nimmo, M. and Shin, Y. (2023). Recent developments of the autoregressive distributed lag modelling framework, *Journal of Economic Surveys*, **37**(1), 7–32.
- de Paula, A., Rasul, I. and Souza, P. (2024). Identifying network ties from panel data: Theory and an application to tax competition, *Review of Economic Studies*, <https://doi.org/10.1093/restud/rdae088>.
- Debarys, N. and Le Gallo, J. (2024). The empirical content of spatial spillovers: Identification issues, SSRN, <http://dx.doi.org/10.2139/ssrn.4751335>.
- Debarys, N. and LeSage, J. P. (2022). Bayesian model averaging for spatial autoregressive models based on convex combinations of different types of connectivity matrices, *Journal of Business & Economic Statistics*, **40**(2), 547–558.

- Delgado, M. S. and Florax, R. J. (2015). Difference-in-differences techniques for spatial data: Local autocorrelation and spatial interaction, *Economics Letters*, **137**, 123–126.
- Elhorst, J. P. (2022). The dynamic general nesting spatial econometric model for spatial panels with common factors: Further raising the bar, *Review of Regional Research*, **42**(3), 249–267.
- Elhorst, J. P. (2024). Raising the bar in spatial economic analysis: Two laws of spatial economic modelling, *Spatial Economic Analysis*, **19**(2), 115–132.
- Elhorst, J. P., Gross, M. and Tereanu, E. (2021). Cross-sectional dependence and spillovers in space and time: Where spatial econometrics and global VAR models meet, *Journal of Economic Surveys*, **35**(1), 192–226.
- Ermagun, A. and Levinson, D. (2018). Spatiotemporal traffic forecasting: review and proposed directions, *Transport Reviews*, **38**(6), 786–814.
- Fruehwirth, J. C. (2013). Identifying peer achievement spillovers: Implications for desegregation and the achievement gap, *Quantitative Economics*, **4**(1), 85–124.
- Gao, H. and Bradley, J. R. (2019). Bayesian analysis of areal data with unknown adjacencies using the stochastic edge mixed effects model, *Spatial Statistics*, **31**, <https://doi.org/10.1016/j.spasta.2019.100357>.
- Geniaux, G. and Martinetti, D. (2018). A new method for dealing simultaneously with spatial autocorrelation and spatial heterogeneity in regression models, *Regional Science and Urban Economics*, **72**, 74–85.
- Gibbons, S. and Overman, H. G. (2012). Mostly pointless spatial econometrics?, *Journal of Regional Science*, **52**(2), 172–191.
- Greenaway-McGrevy, R. and Phillips, P. C. (2023). The impact of upzoning on housing construction in Auckland, *Journal of Urban Economics*, **136**, <https://doi.org/10.1016/j.jue.2023.103555>.
- Griffith, D. A. (2020). Some guidelines for specifying the geographic weights matrix contained in spatial statistical models, *Practical Handbook of Spatial Statistics* (ed. S. Arlinghaus), CRC Press, Boca Raton.
- Halleck Vega, S. and Elhorst, J. P. (2015). The SLX model, *Journal of Regional Science*, **55**(3), 339–363.
- 原祐輔, 花岡洋平, 桑原雅夫 (2016). 道路ネットワーク内の関係性に着目した長期観測プローブデータによるプローブ未観測リンクの交通状態補間, *交通工学論文集*, **2**(1), 1–10.
- Hara, Y., Suzuki, J., and Kuwahara, M. (2018). Network-wide traffic state estimation using a mixture Gaussian graphical model and graphical lasso, *Transportation Research Part C*, **86**, 622–638.
- Holmes, T. J. (1998). The effect of state policies on the location of manufacturing: Evidence from state borders, *Journal of Political Economy*, **106**(4), 667–705.
- Hoshino, T. (2019). Two-step estimation of incomplete information social interaction models with sample selection, *Journal of Business & Economic Statistics*, **37**(4), 598–612.
- Hoshino, T. (2022). Sieve IV estimation of cross-sectional interaction models with nonparametric endogenous effect, *Journal of Econometrics*, **229**(2), 263–275.
- Hudgens, M. G. and Halloran, M. E. (2008). Toward causal inference with interference, *Journal of the American Statistical Association*, **103**(482), 832–842.
- Imbens, G. and Rubin, D. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press, Cambridge.
- Kelejian, H.H. and Piras, G. (2017). *Spatial Econometrics*, 1st ed., Academic Press, Cambridge.
- Kelejian, H.H. and Prucha, I.R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Economics*, **17**(1), 99–121.
- King, G., Honaker, J., Joseph, A. and Scheve, K. (2001). Analyzing incomplete political science data:

- An alternative algorithm for multiple imputation, *American Political Science Review*, **95**(1), 49–69.
- Klein, A. and Tchuente, G. (2023). Spatial differencing for sample selection models with ‘site-specific’ unobserved local effects, *The Econometrics Journal*, **27**(2), 235–257.
- Kolak, M. and Anselin, L. (2020). A spatial perspective on the econometrics of program evaluation, *International Regional Science Review*, **43**(1-2), 128–153.
- Krisztin, T. and Piribauer, P. (2023). A Bayesian approach for the estimation of weight matrices in spatial autoregressive models, *Spatial Economic Analysis*, **18**(1), 44–63.
- Kuminoff, N. V., Parmeter, C. F. and Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?, *Journal of Environmental Economics and Management*, **60**(3), 145–160.
- Lam, C. and Souza, P. C. (2020). Estimation and selection of spatial weight matrix in a spatial lag model, *Journal of Business & Economic Statistics*, **38**(3), 693–710.
- Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models, *Econometrica*, **72**(6), 1899–1925.
- Lee, L. F. (2023). *Spatial Econometrics: Spatial Autoregressive Models*, World Scientific Publishing, Singapore.
- LeSage, J. P. and Chih, Y. Y. (2018). A Bayesian spatial panel model with heterogeneous coefficients, *Regional Science and Urban Economics*, **72**, 58–73.
- LeSage, J. P. and Pace, R. K. (2004). Models for spatially dependent missing data, *The Journal of Real Estate Finance and Economics*, **29**, 233–254.
- LeSage, J. P. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*, Chapman & Hall/CRC, Boca Raton. (各務和彦, 和合肇 訳 (2020). 『入門空間計量経済学』, 彩流社, 東京.)
- Lyytikäinen, T. (2012). Tax competition among local governments: Evidence from a property tax reform in Finland, *Journal of Public Economics*, **96**(7-8), 584–595.
- Mínguez, R., Basile, R. and Durbán, M. (2022). An introduction to pspatreg: A new R package for semiparametric spatial autoregressive analysis, *Region*, **9**(2), R1–R15.
- Mur, J. (2013). Causality, uncertainty and identification: Three issues on the spatial econometrics agenda, *Scienze Regionali*, **12**(1), 5–27.
- 村上大輔 (2024). 『Rではじめる地理空間データの統計解析入門』, 講談社, 東京.
- 奥井亮 (2014). 因子モデルに関する近年の計量経済学研究の進展, 日本統計学会誌, **43**(2), 247–273.
- Olmo, J. and Sanso-Navarro, M. (2023). A nonparametric spatial regression model using partitioning estimators, *Econometrics and Statistics*, <https://doi.org/10.1016/j.ecosta.2023.02.003>.
- Otto, P., Doğan, O. and Taşpınar, S. (2023). A dynamic spatiotemporal stochastic volatility model with an application to environmental risks, *Econometrics and Statistics*, <https://doi.org/10.1016/j.ecosta.2023.11.002>.
- Otto, P., Doğan, O. and Taşpınar, S. (2024a). Dynamic spatiotemporal ARCH models, *Spatial Economic Analysis*, **19**(2), 250–271.
- Otto, P., Doğan, O., Taşpınar, S., Schmid, W. and Bera, A. K. (2024b). Spatial and spatiotemporal volatility models: A review, *Journal of Economic Surveys*, <https://doi.org/10.1111/joes.12643>.
- Paelinck, J. and Klaassen, L. (1979). *Spatial Econometrics*, Saxon House, Farnborough.
- Pearl, J. and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*, Basic Books, New York. (夏目大 訳 (2022). 『因果推論の科学「なぜ？」の問いにどう答えるか』, 文藝春秋, 東京.)
- Pesaran, M. H. (2015). *Time Series and Panel Data Econometrics*, Oxford University Press, Oxford.
- Pinkse, J. and Slade, M. E. (2010). The future of spatial econometrics, *Journal of Regional Science*, **50**(1), 103–117.

- Pinkse, J., Slade, M. E., and Brett, C. (2002). Spatial price competition: A semiparametric approach, *Econometrica*, **70**(3), 1111–1153.
- Piribauer, P., Glocker, C. and Krisztin, T. (2023). Beyond distance: The spatial relationships of European regional economic growth, *Journal of Economic Dynamics and Control*, **155**, <https://doi.org/10.1016/j.jedc.2023.104735>.
- Qu, X. and Lee, L. F. (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix, *Journal of Econometrics*, **184**(2), 209–232.
- Qu, X., Lee, L. F. and Yu, J. (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices, *Journal of Econometrics*, **197**(2), 173–201.
- Qu, X., Lee, L. F. and Yang, C. (2021). Estimation of a SAR model with endogenous spatial weights constructed by bilateral variables, *Journal of Econometrics*, **221**(1), 180–197.
- Reich, B. J., Yang, S., Guan, Y., Giffin, A. B., Miller, M. J. and Rappold, A. (2021). A review of spatial causal inference methods for environmental and epidemiological applications, *International Statistical Review*, **89**(3), 605–634.
- Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment, *Journal of the American Statistical Association*, **75**(371), 591–593.
- Santi, F., Dickson, M. M., Espa, G., Taufer, E. and Mazzitelli, A. (2021a). Handling spatial dependence under unknown unit locations, *Spatial Economic Analysis*, **16**(2), 194–216.
- Santi, F., Dickson, M. M., Giuliani, D., Arbia, G. and Espa, G. (2021b). Reduced-bias estimation of spatial autoregressive models with incompletely geocoded data, *Computational Statistics*, **36**(4), 2563–2590.
- 瀬谷創, 堤盛人 (2014). 『空間統計学：自然科学から人文・社会科学まで』, 朝倉書店, 東京.
- Seya, H., Tsutsumi, M. and Yamagata, Y. (2012). Income convergence in Japan: A Bayesian spatial Durbin model approach, *Economic Modelling*, **29**(1), 60–71.
- Seya, H., Yamagata, Y. and Tsutsumi, M. (2013). Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach, *Regional Science and Urban Economics*, **43**(3), 429–444.
- Seya, H., Tomari, M. and Uno, S. (2021). Parameter estimation in spatial econometric models with non-random missing data, *Applied Economics Letters*, **28**(6), 440–446.
- Shi, W. and Lee, L. F. (2017). Spatial dynamic panel data models with interactive fixed effects, *Journal of Econometrics*, **197**(2), 323–347.
- Shirota, S. and Gelfand, A. E. (2022). Preferential sampling for bivariate spatial data, *Spatial Statistics*, **51**, <https://doi.org/10.1016/j.spasta.2022.100674>.
- Stakhovych, S. and Bijmolt, T. H. (2009). Specification of spatial models: A simulation study on weights matrices, *Papers in Regional Science*, **88**(2), 389–409.
- Sun, Y. (2016). Functional-coefficient spatial autoregressive models with nonparametric spatial weights, *Journal of Econometrics*, **195**(1), 134–153.
- Sun, Y. (2024). Semiparametric spatial autoregressive models with nonlinear endogeneity, *Econometric Reviews*, **43**(6), 434–451, <https://doi.org/10.1080/07474938.2024.2339149>.
- 塚井誠人, 小林潔司 (2007). 長期記憶性を考慮した社会資本の生産性測定, 土木学会論文集 D, **63**(3), 255–274.
- 塚井誠人, 奥村誠 (2007). 空間データセットにおける内生的な空間相関構造の同定に関する考察, 土木計画学研究・講演集, **35**, http://library.jsce.or.jp/jsce/open/00039/200706_no35/pdf/119.pdf (最終アクセス日 2024年9月12日).
- 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**, 357–380.

- 堤盛人, 瀬谷創 (2012a). 応用空間統計学の二つの潮流：空間統計学と空間計量経済学, *統計数理*, **60**(1), 3–25.
- 堤盛人, 瀬谷創 (2012b). 土木計画における応用空間統計学の可能性, *土木学会論文集 D3(土木計画学)*, **68**(5), I_1–I_20.
- von Graevenitz, K. and Panduro, T. E. (2015). An alternative to the standard spatial econometric approaches in hedonic house price models, *Land Economics*, **91**(2), 386–409.
- Wang, W. and Lee, L. F. (2013). Estimation of spatial autoregressive models with randomly missing data in the dependent variable, *The Econometrics Journal*, **16**(1), 73–102.
- Wang, Z. and Song, Y. (2023). Deep learning for the spatial additive autoregressive model with nonparametric endogenous effect, *Spatial Statistics*, **55**, <https://doi.org/10.1016/j.spasta.2023.100743>.
- Wu, J. and Matsuda, Y. (2021). A threshold extension of spatial dynamic panel model with fixed effects, *Journal of Spatial Econometrics*, **2**(3), <https://doi.org/10.1007/s43071-021-00008-1>.
- Xiao, S., Song, Y. and Wang, Z. (2023). Nonparametric spatial autoregressive model using deep neural networks, *Spatial Statistics*, **57**, <https://doi.org/10.1016/j.spasta.2023.100766>.
- Yang, K. and Lee, L. F. (2019). Identification and estimation of spatial dynamic panel simultaneous equations models, *Regional Science and Urban Economics*, **76**, 32–46.
- Yang, Z., Hong, Z., Zhou, R. and Ai, H. (2022). Graph convolutional network-based model for megacity real estate valuation, *IEEE Access*, **10**, 104811–104828.
- 要藤正任, 吉村有博 (2016). 社会資本によるスピルオーバー効果と地域経済成長：市町村データを用いた高速道路整備効果の実証分析, KIER Discussion Paper, No.1603, Institute of Economic Research, Kyoto University, 京都.
- Zeng, H., Zhong, W. and Xu, X. (2024). Transfer learning for spatial autoregressive models, arXiv preprint, <https://doi.org/10.48550/arXiv.2405.15600>.
- Zhu, D., Liu, Y., Yao, X. and Fischer, M. M. (2022). Spatial regression graph convolutional neural networks: A deep learning paradigm for spatial multivariate distributions, *GeoInformatica*, **26**(4), 645–676.

Recent Methodological Developments in Spatial Econometrics

Hajime Seya and Masashi Tomari

Department of Civil Engineering, Graduate School of Engineering, Kobe University

Almost 50 years have passed since the emergence of the field of spatial econometrics. Standard methodologies were largely established by the 2000s, and a vast amount of empirical research has been conducted to date. This paper discusses our personal views on recent methodological developments in the field of spatial econometrics. Specifically, we review important recent methodological developments in terms of 1) identification and causal inference, 2) specification of spatial weight matrices, 3) flexible model structures, 4) spatiotemporal data modeling, and 5) dirty data modeling.