ラプラシアン行列に基づく空間可変係数モデルの提案と街路単位の犯罪分析への応用

許 芮溧*・井上 亮**

Graph Laplacian-Based Spatially Varying Coefficient Model and Its Application to Street-Segment Crime Analysis

Ruili XU*, Ryo INOUE**

Abstract: Traditional spatial regression often relies on Euclidean proximity and analyses aggregated to grid or administrative units, which can overlook network-driven, link-level heterogeneity. Advances in sensing technology and open data collection have produced diverse, high-resolution spatial datasets, creating a demand for regression methods that accommodate heterogeneous, multi-scale structures. Moreover, many observations are inherently based on networks—e.g., crime concentrated along street segments or passenger flows on transit links—yet methods that explicitly model link or node level processes remain underdeveloped. Therefore, in this study, we propose a framework that leverages graph Laplacian eigenvectors as multiscale basis functions to model spatially varying coefficients (SVC) on road networks. By combining eigenvector spatial filtering with Laplacian spectra, our approach enables inference at the street segment scale, and better captures localized network heterogeneity. At last, we demonstrate feasibility and empirical utility using crime calls for service data in New York City.

Keywords: 空間的異質性（spatial heterogeneity），空間可変係数（spatially varying coefficients），ネットワーク（network），グラフラプラシアン（graph Laplacian）

1. はじめに

センサ技術の発達やオープンデータ施策の進展に伴い，空間データの多様化・大規模化は著しく，空間解像度の高い詳細な空間データが入手可能になり，それに応じた分布・精度・構造の異なる幅広い時空間データのための回帰法が求められている．

空間データの基本特性である空間的異質性を考慮した解析の理論や方法は，空間統計学を中心に発展を遂げてきた．空間的異質性は空間現象の生成要因・過程が場所によって異なるという性質であり，従来の分析手法として，空間全域で連続的な変化と特定の地域で不連続的な変化の二種類の仮定が提案される．前者は，地理的カーネルを用いて地理的局所的な分析に基づく地理的加重回帰法 (Fotheringham et al., 1998) や，位置によって生成要因の影響が徐々に変化する空間構造を捉える空間可変係数モデル (Spatially Varying Coefficient, SVC) が代表モデルとして挙げられる．これらの一例として，空間フィルタリング (Eigenvector Spatial Filtering, ESF) を応用したSVC モデル(ESF-SVC) (Griffith, 2008) が存在し，空間的な近接性を記述する行列の固有ベクトルの線形和をあてはめることで空間上での連続的な変化を推定しようというものである．一方，後者は最小分析空間単位とする地域ごとに各地域に異なる係数を設定したモデルを用い，スパースモデリング手法を活用して不連続に変化する空間構造を捉える分析が検討されてきた．

空間的異質性を統計学的に検証する際，モデルの選択を問わず，空間重み行列の設定が重要である．空間重み行列は空間オブジェクト間の隣接関係や距離をもとに，「近さ」を表現するための指標である．従来，地点間のユークリッド距離を元に地理的近接性を定義した分析が行われることが多い．しかし，高解像度の地理分析を行う場合，特に都市内の社会経済現象を対象とする際は，現象が道路網などのネットワーク構造の影響を受けて生起していることを踏まえる必要性が高い．また，分析対象の現象の観測は，道路区間単位や交差点単位での交通事故件数など，リンクやノード単位に行われていることもある．また，現象の影響要因も，沿道道路環境や交差点構造などリンクやノードの性質であることも多い．これらを考えた分析を行うためには，ネットワーク構造に基づく空間分析が不可欠である．

本研究では，ネットワーク構造を表す方法としてラプラシアン行列に注目し，空間的異質性分析への応用を検討する．ラプラシアン行列は，グラフの連結の強さの定量的評価として使われ，固有ベクトルの集合はグラフ分

* 学生会員 東北大学 大学院情報科学研究科（Tohoku University）
〒980-8579 宮城県仙台市青葉区荒巻字青葉 6-6-06 E-mail：xu.ruili.r5@dc.tohoku.ac.jp
** 正会員 東北大学 大学院情報科学研究科（Tohoku University）

割のために重要な情報を提供する(斎藤, 2015)．さらに，ラプラシアン固有ベクトルは大域から局所までの変動を表す基底になり得る(Belkin & Niyogi, 2003)．したがって，ラプラシアン固有ベクトルを説明変数の係数関数の基底として用いることで，異なる空間スケールにおける係数の変動を表現でき，ESF-SVC と同様に空間的可変係数モデルを構成することが可能である．

そこで，本研究は，このラプラシアン行列の固有ベクトルを用いた空間可変係数モデル (Laplacian-SVC) を提案する．道路網などのグラフの隣接関係や重みを考慮した SVC モデルによって，「線 (edge) 単位」「街路セグメント単位」など，より微視的・局所的な空間単位での分析が可能となる．

以降，第 2 章で提案した融合モデルの実行可能性とその推定方法を示し，第 3 章では実データを用いて提案モデルの有用性を示し考察する．最後に，第 4 章において本研究の結論を示す．

２．分析手法

２.１　ESF-SVC モデル

ESF-SVC は，空間的自己相関の代表的な検定統計量である Moran's I 統計量 (MC) に基づいた空間モデリングアプローチである (Griffith, 2008)．n 地点で得られた空間データ$\boldsymbol{z}$ ($n \times 1$ベクトル) の空間相関を表す尺度 MC は，$n \times n$の空間隣接行列$\boldsymbol{C}$，単位行列$\boldsymbol{I}$を用いると，

$$MC = \frac{\boldsymbol{z'MCMz}}{\boldsymbol{z'Mz}} \tag{1}$$

と定式化できる．なお，$\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{11'}/N$，すべての要素が 1 の$n \times 1$のベクトルを$\boldsymbol{1}$と表す．$\boldsymbol{MCM}$の各固有ベクトル$\boldsymbol{e}_l$は空間隣接行列$\boldsymbol{C}$の有する空間構造を表し，$\boldsymbol{e}_l$に対応する固有値$\lambda_l$は空間的自己相関の大きさを表す．ESF-SVC は，この固有ベクトルを説明変数のパラメータβ_kに導入して，β_kが有する連続的空間的異質性を表すモデルで

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{x}_k \circ \boldsymbol{\beta}_k^{ESF} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$
$$\boldsymbol{\beta}_k^{ESF} = \beta_k \boldsymbol{1} + \boldsymbol{E\gamma}_k \tag{2}$$

で表される．ここで，"∘"は要素ごとの積，$\boldsymbol{\beta}_k^{ESF}$は$k$番目の説明変数$\boldsymbol{x}_k$の空間可変パラメータベクトルである．定数$\beta_k$と各固有ベクトルにかかるパラメータベクトル$\boldsymbol{E\gamma}_k$は，それぞれ$\boldsymbol{\beta}_k^{ESF}$の平均値と空間的変動を示す．

２.２　提案手法

ネットワークをノードとそれぞれつなぐリンクで表現し，その隣接関係を隣接行列で記述する．ノード総数をnとすると，隣接行列$\boldsymbol{A}$は，$n \times n$の正方行列で，その成分A_{ij}は，ノードi, jの隣接関係を表し，リンク$\boldsymbol{E}$で結ばれているときに 1，結ばれていないときに 0 とする．

$$A_{ij} = \begin{cases} 1, (i,j) \in \boldsymbol{E} \\ 0, otherwise \end{cases} \tag{3}$$

各ノードに接続するリンクの総数を次数と呼び，次数を対角成分とする行列を次数行列$\boldsymbol{D}$と呼ぶ．その成分D_{ii}は$D_{ii} = \sum_{j=1}^{n} A_{ij} = \sum_{j=1}^{n} A_{ji}$である．

グラフ理論におけるグラフの行列表現であるラプラシアン行列は，$L = D - A$と定義される．この行列は定義から明らかに対称行列である．また，2 次形式は

$$\boldsymbol{x}^T \boldsymbol{Lx} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} \left(x_i - x_j\right)^2 \geq 0 \tag{4}$$

と書けるため，ラプラシアン行列は半正定値であり，固有値は必ず非負で，しかも固有ベクトルはお互いに直交することが知られているため，ラプラシアン行列は以下の形でスペクトル分解できる．

$$\boldsymbol{L} = \sum_{i=1}^{n} \lambda_i \left(\boldsymbol{a}_i \boldsymbol{a}_i^T\right) \tag{5}$$

ここで，λ_iは固有値でて，$\lambda_1 = 0 \leq \lambda_2 \leq \cdots \leq \lambda_n$を満たし，ベクトル$\boldsymbol{a}_i$はそれに対応する単位の固有ベクトルである．ラプラシアン行列のスペクトル解析より，最小固有値λ_1は必ずゼロとなり，最小の正の固有値はグラフの代数的連結度 (algebraic connectivity) と呼ばれ，グラフの連結の強さの定量的評価として使われている．さらに，$\boldsymbol{Lv}_i = \lambda_i \boldsymbol{v}_i$によって得られる固有値$\lambda_i$と固有ベクトル$\boldsymbol{v}_i$はグラフ上の空間モード（周波数）を表せる．すなわち，λ_iが小さいほど，その対応する$\boldsymbol{v}_i$はグラフ上で滑らか（大域に緩やかに変化）なモードを表し，逆に，λ_iが大きいほど，$\boldsymbol{v}_i$は局所的に急変するモードを表す(Belkin & Niyogi, 2003)．

更に，Yamada (2021) は，空間的自己相関測度指標の一つである Geary's C を用い，新たな視点からグラフラプラシアン固有ベクトルを用いた空間パターン表現を議論した．Geary's C は

$$C = \frac{(n-1)}{2\boldsymbol{W}} \frac{\sum_i \sum_j w_{ij} \left(x_i - x_j\right)^2}{\sum_i (x_i - \bar{x})^2}, \boldsymbol{W} = \sum_i \sum_j w_{ij} \tag{6}$$

で表され，式(4)によりラプラシアンを用いて

$$C = \frac{(n-1)}{\boldsymbol{W}} \frac{\boldsymbol{x}^T \boldsymbol{Lx}}{\boldsymbol{x}^T \boldsymbol{Mx}} \tag{7}$$

と書き換えられる．すなわち，Geary's C とラプラシアン固有分解を結びつけると，ラプラシアン固有値は空間的自己相関のスケールを解釈可能となり，ESF へ拡張できる．そこで，ESF-SVC に基づき，グラフ構造を考慮できるラプラシアン行列に基づいた SVC モデルを提案する．

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{x}_k \circ (\beta_k \boldsymbol{1} + \boldsymbol{L\gamma}_k) + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{8}$$

ここで，注意すべき点は，説明変数$\boldsymbol{x}_k$とラプラシアン固有ベクトル$\boldsymbol{L}$との交互項は次元が大きくなりがちで，そのままモデルに組み込むと過学習や計算負荷の増大を招きやすいことである．この問題に対処するため，Generalized Lasso に基づくl_1型の正則化を導入し，交互項の選択を同時に行う必要がある．

３．ニューヨーク市の犯罪通報データへの適用

本章は，ニューヨーク市内の犯罪多発地域を対象に，提案手法を用いて街路区間単位で犯罪通報件数の空間的異質性を分析し，提案手法の有効性を確認する．

犯罪は都市内で特定の場所に偏在する傾向があることが数々の研究で指摘されている．

例えば Weisburd (2015) は，米国の規模の異なる数都市で，街路セグメント単位で犯罪を集計し，都市規模によらず，全犯罪の 50%が 4%程度の街路区間に，25%が 1.5%程度の街路区間に偏在することを示した．

また，近隣物理環境や社会環境，さらには街路の構造的特性が犯罪分布に影響を及ぼす可能性が示唆されている．例えば Gill et al. (2017) は，小規模都市ほど犯罪集中度が高いことを観察した．小規模都市では郊外型の道路構造が多く，大半の道路が一戸建て住宅に面している一方，少数の道路に大型集合住宅が建っており，人口が集中し潜在的な犯罪ターゲットも多いことが影響している可能性があるとしている．また Park (2019) は，平均街路区間長が長く，かつ人口密度が低い地域ほど，犯罪が特定の場所に集中しやすいことを明らかにした．

しかし，都市内で地域や街路ごとに犯罪集中の程度が異なるメカニズムについての理解はまだ十分とは言えない．Weisburd et al. (2024) は，犯罪集中の原因解明が治安政策や犯罪予防の改善に不可欠であることを強調している．例えば，土地利用特性が犯罪集中に影響を与える場合には，土地利用の改変や用途規制を通じて犯罪集中度を緩和し，犯罪発生件数の多い道路の数を減らすことが期待される．そうした施策によって，限られた警察資源をより効果的に配分し，犯罪発生率の高い道路に対する重点的な介入を実現できる可能性があるとしている．従って，犯罪が集中する道路の土地利用特性や道路構造を把握することは，市当局が犯罪防止対策を改善するための有効な手段となり得る．

それらの既往研究を踏まえ，従来手法より柔軟に街路セグメント単位の空間的異質性を捉えることができる提案モデルを用いて，街路構造特性 (幅員や長さ) や近隣土地の土地利用特性が犯罪集中に与える影響を検証する．

３．１　データ

本研究では，ニューヨーク市の calls for service（通報データ）を対象とし，解析単位を街路区間 (street segment) とする．通報データは「通報された全ての事象（確認の有無を問わない)」を包含するため，逮捕記録や事件報告より包括的である．道路ネットワークは DCM (Digital City Map) の道路中心線データを用いる．散歩道の(path)データには幅員などの情報が欠落しているため，道路ネットワークの構築には通り(street)タイプのみを使用した．各街路上の通報件数（集計値）を被説明変数とし，「general crime，violent crime，property crime」をカテゴリ別に集計・分析する (Weisburd at el., 2024)．道路属性（街路区間長と街路幅員）や土地利用当該街路両側に接する土地利用タイプのカウント値（一戸建て・集合住宅・商業用地・空き地）を説明変数とする．また，通報記述に“ambulance”, “train”を含む救急や鉄道関連の記録は非犯罪的通報とみなし分析対象外とした．交差点で発生した通報は，交差点固有の発生機構を持つため街路区間上の事象と性質が異なると判断して除外した．

解析対象を「辺」に置くために，元の道路グラフ $G=(V,E)$ をライン・グラフ$L(G)$に変換した．変換後は元の各辺が$L(G)$の頂点に対応し，元の節点で接続する辺同士が$L(G)$上で隣接する関係となる．これにより「辺上データ」を頂点データとして扱い，ラプラシアン行列を構築し，得られた固有ベクトルを提案モデルに導入して分析する．なお，説明変数は，各パラメータに対する正則化の影響を同一にするため，説明変数の値を標準化した．

３．２　分析結果

本分析では，ニューヨーク市の中で犯罪通報数が最も多い管轄区であるブロンクス区の第 40 管轄区を対象地域として選定した(図 1). 2025 年 1 月から 2 月までの 2 か月間に，本地域では合計 74,499 件の犯罪通報が記録され，そのうち 2,499 件が財産犯であった（図 2）．対象地域の街路区間数は 945 である．ここでは窃盗・強盗など財産犯に関する分析の推定結果のみを示す．図 3 の上段は街路属性に関するパラメータ推定結果，下段は土地利用特性に関するパラメータ推定結果を示す．

街路幅員のパラメータは全域で正かつ一定に推定され，財産犯発生に対して一様な影響を与えており，空間的異

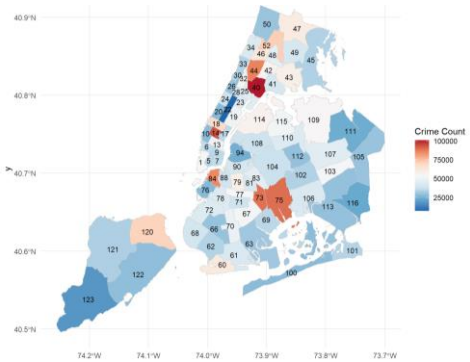


図 1　2025 年ニューヨーク市警察署管轄区ごとの犯罪通報件数



図 2　街路ごとの総通報件数(左)と財産犯(右)の分布

質性は検出されなかった．街路区間長の影響は場所によって正負の効果が現れ，空間的異質性を有することが確認された．特定の街路区間では区間長が延びたら，路上の窃盗や強盗などの財産犯発生が増加する傾向にある．一方，集合住宅件数の影響は正に推定され，集合住宅が増加すると財産犯が増加する傾向が確認された．集合住宅の集積が犯罪発生に影響を与える可能性があるという既往研究の結論を裏付けているが，集合住宅の規模や居住密度が分析に入っていないため，潜在的ターゲットの集中が犯罪発生を助長するのかは判断できない．一戸建て住宅件数のパラメータは負の定数に推定され，財産犯の増加には寄与しないことを示唆する．また，商店数が街路上の財産犯発生に与える影響は地域によって異なり，特に東部の一部街路においては，商店数の増加が財産犯発生の可能性を高める傾向が観察された．商業活動の集中に伴う人の集積や魅力的なターゲットの存在が影響している可能性がある．空き地は，北部および東部に多く存在するが，これらが増加すると一部の街路で財産犯発生が有意に増加する傾向が見られた．

４．おわりに

本研究は，ネットワーク構造の空間的異質性を捉え，より微視的・局所的な空間単位での分析が可能となる手法として，ラプラシアン行列の固有ベクトルを用いた空間可変係数モデルを提案した．

提案手法の有効性を確認するため，ニューヨーク市の犯罪多発地域内の犯罪通報データに適用した．分析結果から，既往研究の結論と一致して，街路区間長と街路両側の集合住宅数の増加が犯罪発生の増加に寄与することが確認できた．さらに，土地利用が街路ごとで犯罪発生への影響は，街路区間ごとで異なることが分かった．

今後は，犯罪のような係数データに適したポアソンモデルと提案モデルを統合し，発生頻度のモデリング精度を高める予定である．併せて，大規模ネットワークへの適用性と計算効率の改善を図るとともに，ユークリッド距離に基づく従来手法との比較を行い，ネットワーク解析の実用的有用性と限界を改めて検証する．

参考文献

斎藤直樹. (2015). グラフ・ネットワーク上での応用調和解析. 応用数理, 25(3), 102-111.

Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, *15*(6), 1373-1396.

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30(11), 1905-1927.

Gill, C., Wooditch, A., & Weisburd, D. (2017). Testing the “law of crime concentration at place” in a suburban setting: Implications for research and practice. *Journal of Quantitative Criminology*, 33(3), 519-545.

Griffith, D. A. (2008). Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environment and Planning A*, 40(11), 2751-2769.

Park, S. (2019). *Examining the “Law of Crime Concentrations” across Multiple Jurisdictions* (Doctoral dissertation, George Mason University).

Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology*, 53(2), 133-157.

Weisburd, D., Zastrow, T., Kuen, K., & Andresen, M. A. (2024). Crime concentrations at micro places: A review of the evidence. *Aggression and Violent Behavior*, 78, 101979.

Yamada, H. (2021). Geary’s C and spectral graph theory. *Mathematics*, 9(19), 2465.

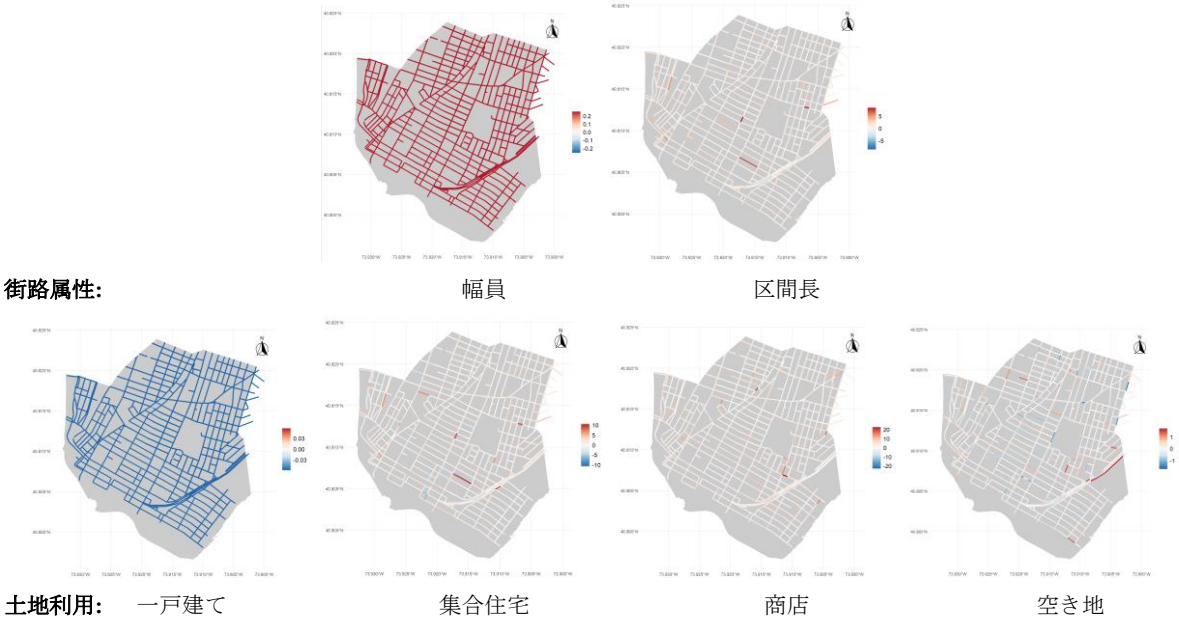


図３　財産に関する犯罪に対するパラメータの推定結果