

異なる空間スケールの異質性を考慮可能な分析法の提案 空間可変パラメータモデルとスパースモデリングの融合アプローチ 田 皓一郎*・井上 亮**

Extraction of Multi-Scale Spatial Heterogeneities Fusion Approach of Spatially Varying Coefficient Model and Sparse Modeling

Koichiro DEN*, Ryo Inoue**

Spatial data have various spatial heterogeneity that is different in scale. For example, in a real estate market in an urban area, the difference of valuation at between city center and suburb causes large-scale spatial heterogeneity, and the existence of upmarket areas causes local-scale spatial heterogeneity. Nevertheless, few studies have considered those different scales of spatial heterogeneity simultaneously. We propose a model that combines an Eigenvector Spatial Filtering based Spatial Varying Coefficient model (ESF-SVC) and fused lasso to analyze that have spatial heterogeneity with different scales. We evaluated the proposed method through the application to simulation data and real estate rent data.

Keywords: 空間的異質性 (spatial heterogeneity), 空間可変パラメータモデル (spatial varying coefficient model), スパースモデリング (sparse modeling), 空間スケール (spatial scale)

1. はじめに

近年、行政機関や民間企業の積極的なデータ公開によって、空間解像度の高い詳細な空間データが入手可能になり、社会経済活動の実態を、より詳細に、定量的に把握できるようになってきた。

空間データの分析を行うアプローチのひとつとして、説明変数と被説明変数の関係を定量的に分析する回帰モデルがあげられる。最も基本的な線形回帰モデルは、被説明変数 y と説明変数 X 、誤差項ベクトル ϵ を用いて、

$$y = X\beta + \epsilon \quad \epsilon \sim (0, \sigma^2 I) \quad (1)$$

で与えられる。しかし、被説明変数 y と説明変数 X との関係が場所によって異なる、空間的異質性を有する現象を線形回帰モデルで分析すると、誤差項の分散は場所毎に異なり、また、誤差項は空間的自己相関を持つため、誤差項に対して仮定した分散均一・共分散なしの条件を満たされないため、推定パラメータはバイアスを有する。

空間的異質性を有する現象分析に対処するため、パラメータ β が地域、あるいは、地点ごとに異なるモデルが提案されてきた。このようなモデルは、推計

されたパラメータの空間パターンからデータの有する空間的異質性の実態把握ができ、地理的な解釈が行える点で、有用であると言える。空間的異質性を考慮した研究として、地理的加重回帰 (Brunsdon et al., 1996) に代表される空間可変パラメータ (Spatially Varying Coefficient: SVC) モデルが挙げられる。これらのモデルで推定されたパラメータの空間分布を観察することで、現象が有する空間的異質性の地理的な特性を捉える分析が行われている (例えば, Lu et al., 2014; 崔・鈴木, 2011)。また、対象地域をいくつかのセグメントに分割をして、セグメントごとのパラメータ推計を行うことで、空間的異質性を考慮する試みも行われている。例えば、小学校区を空間単位の基準に分割し、分析をした Goodman and Thibodeau (1998) や、町丁目空間単位の基準に分割し、分析した Inoue et al. (2020) などがある。

前者では、隣接地域と連続的に変化する大域的な異質性が、後者では、隣接地域と不連続に変化する局所的な異質性が検出される。しかし、例えば不動産価格は、都心から郊外にかけて、物件の広さや交通利便性などの属性に対する評価が徐々に変化する

* 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 : koichiro.den.r2@dc.tohoku.ac.jp

** 正会員 東北大学 大学院情報科学研究科 (Tohoku University)

大域的な異質性と、ブランド地名・駅名で評価が高い町丁目や駅勢圏の境界で価格が不連続に変化する局所的な異質性の、両方が存在する。しかし、既往研究では、これら2種類の異なる空間的異質性を同時に考慮し、両者を区別して検出する方法について議論は十分されていない。

そこで本研究では、空間可変パラメータモデルの一種である Eigenvector Spatial Filtering based spatially varying coefficient (ESF-SVC) モデルと、スパースモデリングの一種である Fused LASSO を融合したモデルを用いて、大域的、局所的両方の空間的異質性のパターンを抽出する手法を提案する。

まず、第2章で ESF-SVC と Fused LASSO を概説した後、それらを融合した提案手法を示す。続いて、第3章でシミュレーションデータを用い、提案モデルの有用性を示し、第4章では、不動産価格データを用いた実データ分析の結果を示しながら考察を行う。最後に、第5章で本研究の結論を示す。

2. 融合モデルの提案

2.1. ESF-SVC モデル

空間的自己相関の検定統計量として代表的なものとして Moran's I 統計量 (MC) がある。変数 $\mathbf{z}$ の Moran's I 統計量は、 n 地点間の空間近接性行列 $\mathbf{C}$ と、 n 個の1からなるベクトル $\mathbf{1}$ 、単位行列 $\mathbf{I}$ を用いて、

$$MC = \frac{\mathbf{z}'(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{C}(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{z}}{\mathbf{z}'(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{z}} \quad (2)$$

で与えられる。式(2)の $(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{C}(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)$ の各固有ベクトルは空間近接性行列 $\mathbf{C}$ の有する空間相関構造を表現している。

Griffith (1996) は、行列 $(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{C}(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)$ の固有ベクトルを線形回帰モデルの説明変数として導入して空間相関を考慮する、Eigenvector Spatial Filtering (ESF) を提案した。ESF の基本式は、 k 列目が行列 $(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{C}(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)$ の第 k 固有ベクトルである行列 $\mathbf{E}$ を用いて、

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim (\mathbf{0}, \sigma^2\mathbf{I}) \quad (3)$$

で与えられる。

ESF-SVC (ESF-based spatially varying coefficient) モデル(Griffith, 2008)は、説明変数 $\mathbf{x}_k$ のパラメータ β_k の空間的異質性を考慮するため、 $\beta_k^{\text{ESF}} = \beta_k \mathbf{1} + \mathbf{E}\boldsymbol{\gamma}_k$ と

し、地点ごとに異なるパラメータを推定する。そのモデルは以下で与えられる。

$$\mathbf{y} = \sum_{k=1}^K \mathbf{x}_k \beta_k^{\text{ESF}} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim (\mathbf{0}, \sigma^2\mathbf{I}) \quad (4)$$

2.2. Fused LASSO

Fused LASSO (Tibshirani et al., 2004) は、パラメータ推定の目的関数に罰則項を加えて変数選択を行うスパースモデリングの1つである。

最も代表的なスパースモデリング手法は、 L_1 罰則項を導入した LASSO (least absolute shrinkage and selection operator) である (Tibshirani, 1996)。LASSO による線形回帰モデルの推定は、罰則項の重みを表すハイパーパラメータ λ を用いて式(5)で与えられる。

$$\min_{\boldsymbol{\beta}} \frac{1}{2} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{i,j} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (5)$$

Fused LASSO は、パラメータの値に対する L_1 罰則項に加え、隣接パラメータの差に L_1 罰則項を与えたモデルである。隣接点で変化が生じた場合のみ、異なる値のパラメータを推定する。回帰分析における Fused LASSO の推定式は2種類の罰則項に対する重みを定めるハイパーパラメータ λ, δ と、隣接関係の組み合わせの集合 $\mathbf{F}$ を用い、式(6)で与えられる。

$$\min_{\boldsymbol{\beta}} \frac{1}{2} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{i,j} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| + \delta \lambda \sum_{\mathbf{F}} |\beta_j - \beta_l| \quad (6)$$

井上 (2019)、井上ら (2020) は、不動産価格分析において、町丁目ごとに設定したダミー変数のパラメータに対し隣接関係の L_1 罰則項を加え、パラメータが同じ値に推定された町丁目を抽出することで、局所的に価格水準が異なる地域を、町丁目単位に抽出した。このように、近年 Fused LASSO は地理的分析への応用が試みられている。

2.3. 提案モデル

本研究は、大域的な空間的異質性を表現する ESF-SVC モデルと、小地域ごとの局所的な空間的異質性を抽出する Fused LASSO を融合し、様々なスケールの空間的異質性を抽出する手法を提案する。

大域的な空間的異質性は、切片 β_0^{ESF} と説明変数の

パラメータ β_k^{ESF} で表現する．本研究は，Griffith and Chun (2014) 等で提案されている，第 1 固有値に対応する固有ベクトルの Moran's I 統計量 MC_1 と第 j 固有値に対応する固有ベクトルの Moran's I 統計量 MC_j の比が $MC_j/MC_1 > 0.25$ を満たす固有ベクトルを説明変数として利用する．局所的な空間的異質性は，空間を小地域に分割し，各地域に割り当てられたダミー変数のパラメータ β^{area} で表現する．この回帰モデルは，小地域のダミー変数行列 D^{area} を用いて式(7)で表せる．

$$y = \beta_0^{\text{ESF}} + \sum_{k=1}^K X_k \circ \beta_k^{\text{ESF}} + D^{\text{area}} \beta^{\text{area}} + \varepsilon$$

$$\beta_0^{\text{ESF}} = \beta_0 \mathbf{1} + E\gamma_0$$

$$\beta_k^{\text{ESF}} = \beta_k \mathbf{1} + E\gamma_k \quad (7)$$

このモデルに Fused LASSO の正則化項を導入した融合モデルの最適化問題は，隣接する小地域の組み合わせの集合 F ，ハイパーパラメータ $\lambda, \delta_1, \delta_2$ とすると，式(8)で表せる．

$$\min_{\beta, \gamma} \frac{1}{2} \sum_{i=1}^N \left[y_i - \left(\sum_{k=0}^K x_{i,k} \beta_k + \sum_{k=0}^K \sum_{l=1}^L x_{i,k} e_{i,l} \gamma_{l,k} + \sum_{m=1}^M d_{i,m} \beta_m^{\text{area}} \right) \right]^2$$

$$+ \sum_{(p,q) \in F} |\beta_p^{\text{area}} - \beta_q^{\text{area}}|$$

$$+ \delta_1 \lambda \sum_{m=1}^M |\beta_m^{\text{area}}| + \delta_2 \lambda \sum_{k=0}^K \sum_{l=1}^L |\gamma_{l,k}| \quad (8)$$

3. シミュレーション実験による性能評価

本章では，大域的・局所的な空間的異質性を有するシミュレーションデータを生成し，提案手法と ESF-SVC，Fused LASSO によるモデルを推定した結果を比較して，提案モデルの性能を評価する．

3.1. シミュレーション実験の概要

(1) シミュレーションデータの生成

シミュレーションデータにおける観測点は，正方形領域内にランダムに 4,000 点発生させる．各点間の距離行列から $(I - 11'/n)C(I - 11'/n)$ の固有ベクトルを計算し，特定の固有ベクトルパラメータをゼロ以外に設定して，ESF 項で表現可能な大域的な空間的異質性を有するデータを生成する．なお，空間

近接性行列 C の要素は，Gaussian 型のカーネル関数を用いて設定する．

また，領域を小地域に分割し，その地域ごとにダミー変数を付与する．特定の小地域のダミー変数パラメータのみをゼロ以外に設定し，他の地域とは異なる特徴を持つ小地域が存在する，局所的な空間的異質性を有するデータを生成する．

3.2.~3.4.では，大域的・局所的な空間的異質性の両方が存在するデータ，大域的な空間的異質性のみ存在するデータ，局所的な空間的異質性のみ存在するデータに対応するパラメータを設定した上で，平均 0，分散 0.2 の正規分布に従う乱数を加えて，データを生成する．

(2) モデルの性能評価

モデルの推定精度は，残差の Moran's I 統計量，RSS(残差平方和)，AIC(赤池情報量基準)で比較する．また，抽出されたパラメータの空間分布の観察も同時に行い精度の比較を行う．

なお，提案モデルと Fused LASSO は，R の genlasso パッケージによって推定を行った．このパッケージは，隣接パラメータとの差に関する正則化項のハイパーパラメータとパラメータ自体に課す正則化項に対する重みの比 δ を固定した下で， λ を徐々に変化させ，それぞれで得られる変数選択結果が求められるものである．本研究では， $\delta_1 = \{0.1, 1, 10\}$ ， $\delta_2 = \{0.1, 1, 10\}$ ，Fused LASSO の推定では $\delta = \{0.1, 1, 10\}$ と予め設定し，AIC 最小の推定結果を探索する．

3.2. 大域的・局所的両方の空間的異質性が存在するデータでの手法比較

大域的，局所的両方の空間的異質性が存在する， $\beta_0 = 1$ ， $\gamma_{1,0} = 10$ ， $\gamma_{10,0} = 5$ ， $\beta_1 = 2$ ， $\gamma_{5,1} = 5$ ， $\gamma_{15,1} = 15$ ， $\beta_{55}^{\text{area}} = \beta_{56}^{\text{area}} = 1.2$ ， $\beta_{57}^{\text{area}} = 0.6$ の設定で生成したデータを用いた実験結果を表 3.1 に示す．

提案モデルは $\delta_1 = \delta_2 = 1$ ，Fused LASSO は $\delta = 1$ のとき，AIC が最小になった．この場合の各モデルの推定パラメータの空間分布を図 3.1 に示す．図より，ESF-SVC では局所的な空間的異質性が切片で，Fused LASSO では大域的な空間的異質性が小地域ダミーのパラメータで表現され，推定パラメータは，データ生成時の設定と異なることが分かる．一方，

提案モデルでは、大域的な空間的異質性と局所的な空間的異質性を分離して抽出でき、設定時とほぼ同じパラメータが推定できていることが確認できる。加えて、提案手法は RSS, AIC とともに他手法に比べて小さく、説明力と予測性能の観点から優れたモデルと言える。また、残差の Moran's I 統計量は、空間相関がないことを表す 0 に近い値を示しており、提案手法は、シミュレーションデータが有する空間的自己相関を適切にモデルで表現できているといえる。

3.3. 大域的な空間的異質性のみが存在するデータでの手法比較

大域的な空間的異質性のみが存在する、 $\beta_0 = 1$, $\gamma_{1,0} = 10$, $\gamma_{10,0} = 5$, $\beta_1 = 2$, $\gamma_{5,1} = 5$, $\gamma_{15,1} = 15$ の設定で生成したデータを用いた実験結果を表 3.2 に示す。提案モデルは $\{\delta_1, \delta_2\} = \{10, 1\}$, Fused LASSO は $\delta = 1$ のとき、AIC が最小となった。設定パラメータの空間分布と、3 手法の切片・小地域ダミーの推定パラメータの空間分布を図 3.2 に示す。

図より、提案モデルでは、局所的な空間的異質性は抽出されず、大域的な空間的異質性のみ抽出されたことが確認できる。また、提案手法は RSS, AIC とともに ESF-SVC と同程度の値を示し、大域的な空間的異質性のみ存在するデータに対しても高い推定精度を持つことが確認された。

一方、Fused LASSO では AIC が他の 2 手法と比べて著しく大きな値となり、シミュレーションデータが有する大域的な空間的異質性を適切に表現できないことが分かる。

3.4. 局所的な空間的異質性のみが存在するデータでの比較

局所的な空間的異質性のみが存在する、 $\beta_0 = 1$, $\beta_1 = 2$, $\beta_{55}^{area} = \beta_{56}^{area} = 1.2$, $\beta_{57}^{area} = 0.6$ の設定で生成したデータを用いた実験結果を表 3.3 に示す。提案モデルは $\{\delta_1, \delta_2\} = \{1, 10\}$, Fused LASSO は $\delta = 1$ のとき、AIC が最小となった。また、3 手法による推定パラメータの空間分布を図 3.3 に示す。

図より、提案モデルでは、局所的な空間的異質性のみ抽出されたことが確認できる。また、提案手法

は RSS, AIC とともに Fused LASSO と同程度の値を示し、局所的な空間的異質性のみが存在するデータに対しても高い推定精度を持つことが確認された。

一方、ESF-SVC では、局所的な空間的異質性が切片の項に含まれて検出され、AIC が他の 2 手法と比べて大きくなり、局所的な空間的異質性を適切に表現できないことが確認された。

表 3.1 大域・局所的な空間的異質性が存在するデータにおける 3 手法の統計量比較

	$\{\delta_1, \delta_2\}$	δ	AIC	残差の Moran's I	RSS
提案モデル	{1,1}		-1458.7	-0.00329	155.3
	{1,10}		-1284.3	-0.00331	153.78
	{10,1}		-984.13	-0.00183	154.93
	{0.1,0.1}		-845.37	0.00048	175.36
	{10,10}		-1410.5	-0.00443	154.23
	{1,0.1}		-1440	-0.00463	154.29
	{0.1,1}		-723.46	0.0023	172.18
	{0.1,10}		-872.15	-0.00196	173.56
	{10,0.1}		-1300.3	-0.00386	154.6
Fused LASSO		1	-443.43	0.014696	202.62
		10	-546.52	0.065772	259.81
		0.1	-8.8895	0.055276	227.8
ESF-SVC			-0.2893	0.010635	227.15

表 3.2 大域的な空間的異質性のみが存在するデータにおける 3 手法の統計量比較

	$\{\delta_1, \delta_2\}$	δ	AIC	残差の Moran's I	RSS
提案モデル	{10,1}		-1411.7	-0.003578	162.99
Fused LASSO		1	-569.28	0.012981	162.25
ESF-SVC			-1346.2	-0.004271	196.73

表 3.3 局所な空間的異質性のみが存在するデータにおける 3 手法の統計量比較

	$\{\delta_1, \delta_2\}$	δ	AIC	残差の Moran's I	RSS
提案モデル	{1,10}		-1351.6	-0.00314	164.56
Fused LASSO		1	-1351.5	-0.003152	164.65
ESF-SVC			175.71	0.00952	237.37

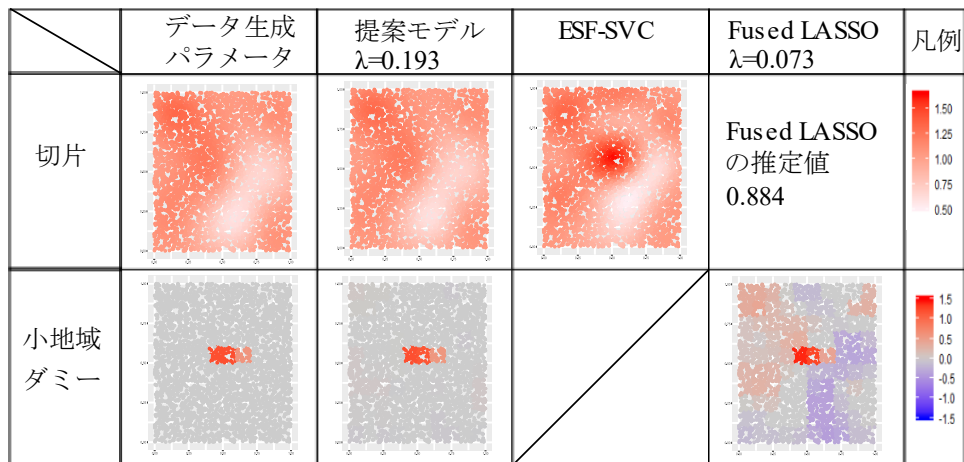


図 3.1 大域的・局所的な空間的異質性が存在するデータに対する
各手法の切片(上), 小地域ダミー(下)の推定パラメータの空間分布

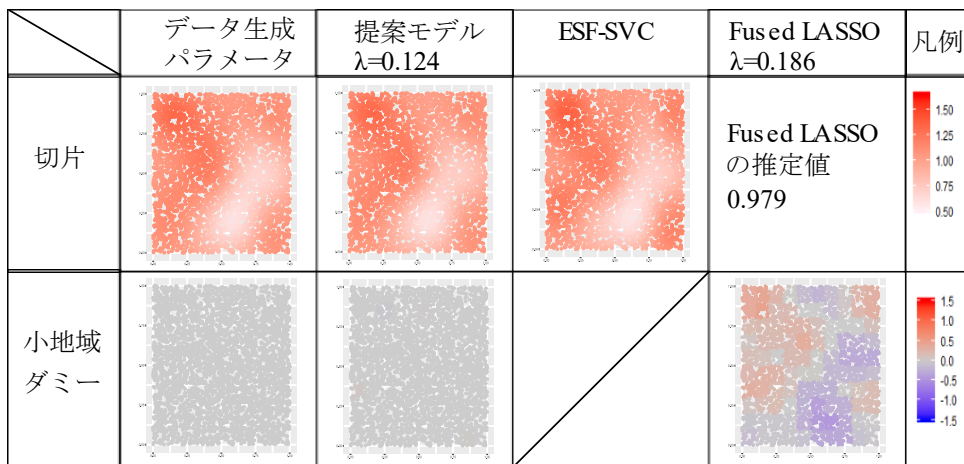


図 3.2 大域的な空間的異質性のみが存在するデータにおける
各手法の切片(上), 小地域ダミー(下)のパラメータ空間分布

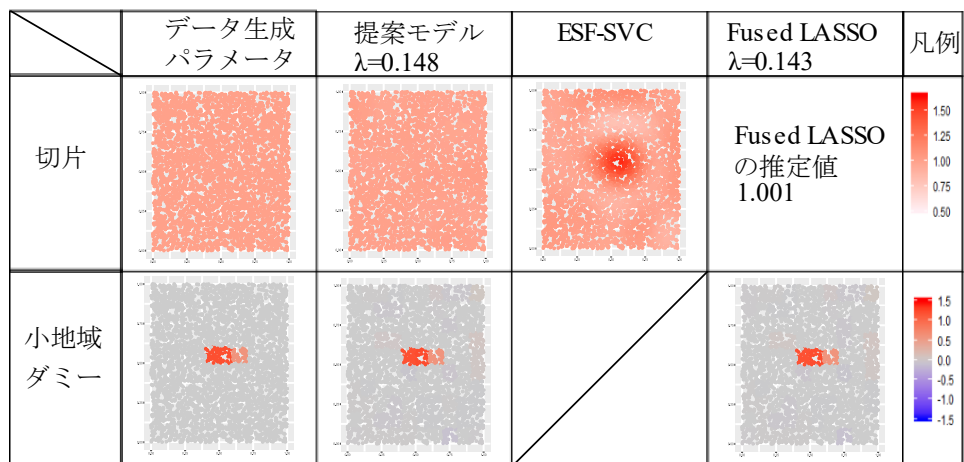


図 3.3 局所的な空間的異質性のみが存在するデータにおける
各手法の切片(上), 小地域ダミー(下)のパラメータ空間分布

3.5. シミュレーション実験のまとめ

本章では、シミュレーションデータを生成し、提案手法と ESF-SVC, Fused LASSO によるモデルを推定した結果を比較して、提案モデルの性能を評価した。分析の結果、提案モデルは大域・局所的な空間的異質性が存在するデータのみならず、大域的な空間的異質性のみ存在するデータ、局所的な空間的異質性のみ存在するデータにおいても、他手法に比べ高い性能で推定できることが確認された。以上より、提案モデルは空間的異質性を考慮するうえで優れたモデルと言える。

4. 実データを用いた検証

4.1. 分析方法

本章では、アットホーム株式会社が 2017 年に収集した渋谷区、世田谷区、目黒区の募集賃料データ 15,263 件を、提案モデル, ESF-SVC, Fused LASSO を用いて分析し、推定結果の残差の Moran's I 統計量, 自由度調整済み決定係数, AIC, 抽出されたパラメータの空間分布を比較する。

被説明変数を単位面積当たり賃料の対数値, 説明変数を, 最寄り駅からの徒歩時間, 物件の専有面積, 築年数, 階層それぞれの対数値と駅勢圏のダミー変数とした。なお, 数値属性の分散の大きさにより, 正則化項がパラメータ推定値に与える影響が変わるため, 数値属性は平均 0, 分散 1 に標準化した。

提案モデル, ESF-SVC における大域的な空間的異質性の抽出には, バンド幅 1km の Gaussian 型カーネル関数に物件間距離を代入して設定した近接性行列 $\mathbf{C}$ を用いて求めた, $(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathbf{C}(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)$ の 19 個の固有ベクトルを説明変数の候補として使用した。

提案モデル, Fused LASSO における局所的な空間的異質性の抽出に使用する地域分割は町丁目単位とした。物件が存在する町丁目にダミー変数を設定し, その隣接関係は, 町丁目ポリゴンが接しているかで判断した。本分析では提案モデルのハイパーパラメータを $\delta_1 = \delta_2 = 1$, Fused LASSO のハイパーパラメータを $\delta = 1$ と固定して λ のみを変化させ, AIC 最小の推定結果を探索した。

表 4.1 3 手法の統計量比較

	AIC	残差の Moran's I	自由度調整済 み決定係数
提案モデル	-20371.6	-0.00005	0.6527
Fused LASSO	-19278.7	-0.00023	0.6271
ESF-SVC	-18626.5	0.00613	0.6155

4.2. 分析結果

ハイパーパラメータ探索の結果, 提案モデルは $\lambda = 0.076$, Fused LASSO は $\lambda = 0.080$ で AIC が最小になった。以下, この推定結果を示す。

表 4.1 に 3 手法の実験結果を示す。まず, 残差の Moran's I に注目すると, いずれも残差から空間相関を除去できていることが確認できる。また, 提案モデルの決定係数は他手法より高く, AIC は他手法より低い。この結果から, 提案モデルは実データに適用した場合においても, 他手法に比べ説明力と予測性能を有すると確認された。

続いて, 図 4.1 に各手法における築年数, 駅からの徒歩時間, 町丁目ダミーのパラメータ推定結果を示す。なお, 町丁目ダミーの黒太線で囲まれた地域は共通のパラメータとして推定された地域を表す。最後に, 提案モデル, Fused LASSO の町丁目ダミーのパラメータ, ESF-SVC における切片の空間分布の, 二子玉川周辺における推定結果を図 4.2 に示す。

図 4.1 の築年数パラメータの推定結果に注目すると, ESF-SVC と Fused LASSO は絶対値や空間変動が小さいのに対し, 提案モデルは西部と東部でパラメータ値が大きく異なることが確認できた。以上より, 提案モデルは築年数の持つ空間的異質性を他の手法よりも大きく評価したといえる。また, 提案モデルのパラメータ空間分布が東部の都心で 0 に近い推定値を取ることから, 都心に近い物件は, 築年数による減価が殆どないと考察できる。

また, 駅からの徒歩時間に対するパラメータに注目すると, 提案モデル, ESF-SVC はともに, 地図中央部と東部で絶対値が小さく, 西部で絶対値が大きく推定された。前者は東急世田谷線, 京王井の頭線の沿線地域, 後者は小田急電鉄小田原線沿線地域である。以上より, 路線によって徒歩時間の評価が異

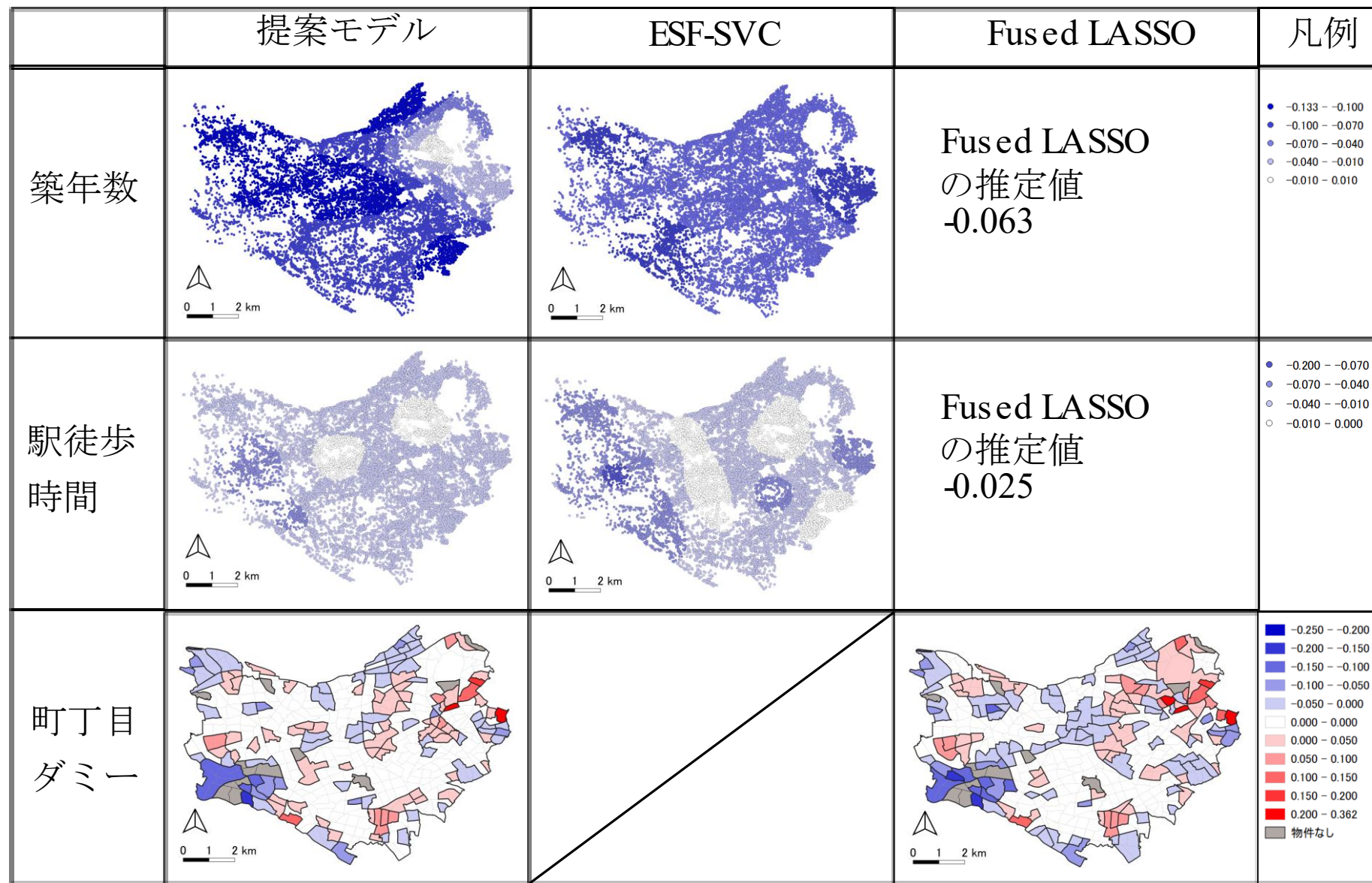


図 4.1 各手法における築年数の空間可変パラメータ，駅からの距離の空間可変パラメータの推定結果，町丁目ダミーパラメータの推定結果

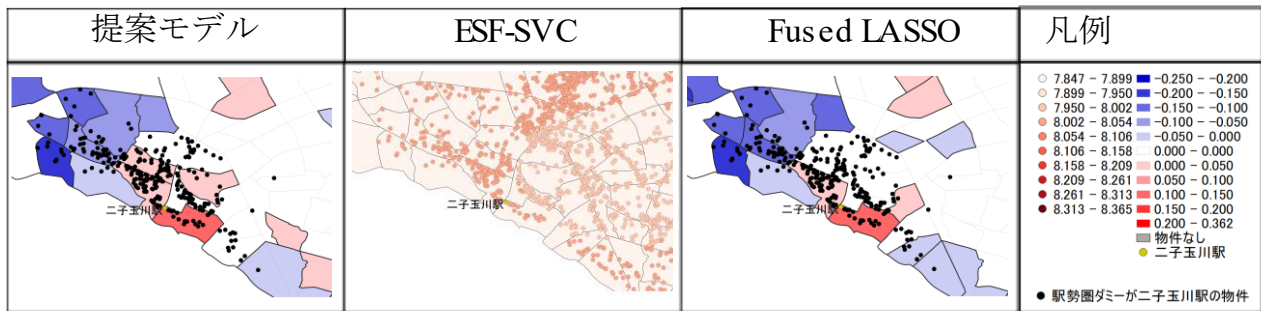


図 4.2 提案モデル・Fused LASSO の町丁目パラメータ，および，ESF-SVC の切片の
二子玉川駅周辺における推定結果

なる可能性が示唆され，特に前者の東急世田谷線に関しては，この路線の駅への近接性は賃料評価に与える影響が小さいことを表していると考えられる。

最後に図 4.2 の提案モデル，Fused LASSO の町丁目パラメータの空間分布より，局所的な空間的異質性は駅勢圏より狭い範囲で存在すると考えられる。また，ESF-SVC における切片推定値の空間分布では，狭い範囲でのパラメータ変動は観測されなかった。よって，ESF-SVC では局所的な空間的異質性のパターンを十分に抽出できないことを確認した。

5. 結論

本研究では，ESF-SVC と Fused LASSO を融合したモデルを用いて，大域的，局所的両方の空間的異質性を抽出する分析手法を提案した。

シミュレーションデータによる分析によって，提案モデルは，大域的・局所的両方の空間的異質性が存在するデータ，大域的な空間的異質性のみ存在するデータ，局所的な空間的異質性のみ存在するデータのいずれに対しても，他手法に比べ高い性能で推定できることが確認された。さらに，提案モデルは実データに適用した場合においても，大域的な空間的異質性の抽出と，町丁目単位の局所的な空間的異質性の抽出が行えることが確認された。

謝辞

本研究は JSPS 科研費 18H01552 の助成を受けた。また，東京大学 CSIS 共同研究 (No. 815) による成果である (利用データ: 不動産データライブラリー戸データ (アットホーム株式会社提供))。

参考文献

- Brunsdon, C., Fotheringham, S., and Charlton M., 1996. Geographically weighted regression: A method for exploring spatial non-stationarity. *Geographical Analysis*, **28**(4): 281-298.
- Goodman, A.C., and Thibodeau, T.G., 1998. Housing market segmentation. *Journal of Housing Economics*, **7**(2): 121-143.
- Griffith, D., 1996. Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying georeferenced data. *The Canadian Geographer*, **40**(4): 351-367
- Griffith, D., 2008. Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environment and Planning A: Economy and Space*, **40**(11): 2751-2769.
- Griffith, D., and Chun, Y., 2014. Spatial autocorrelation and eigenvector spatial filtering. *Handbook of Regional Science*, 1477-1507.
- Inoue, R., Ishiyama, R., and Sugiura, A., 2020. Identifying local differences with fused-MCP: An apartment rental market case study on geographical segmentation detection. *Japanese Journal of Statistics and Data Science*, **3**: 183-214.
- Lu, B., Charlton, M., Harris, P., and Fotheringham, A.S. 2014. Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *International Journal of Geographical Information Science*, **28**(4): 660-681.
- Tibshirani R. 1996. Regression shrinkage and selection via

the lasso, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **58**(1): 267-288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K., 2005. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67: 91-108.

井上 亮. 2019. 局所の特徴を抽出する: Fused-MCPに基づく地域分析の可能性. 第 31 回 RAMP 数理最適化シンポジウム論文集, 143-157.

井上 亮, 石山 里穂子, 杉浦 綾子, 2020. 東京都区部の賃貸マンション市場の地理的分割の実態把握—スパースモデリングによるアプローチ—. 土木学会論文集 D3 (土木計画学), **76**(3): 251-263.

崔 唯爛, 鈴木 勉, 2011. 地理的加重回帰法(GWR)を用いた食料品アクセシビリティの推定~東京都を例に. 地理情報システム学会 学術研究発表大会講演集, **20**, C-6-4.