

氏 名	SEPTIAN RAHARDIANTORO		
授与した学位	博 士		
専攻分野の名称	環境学		
学位授与番号	博乙第	4 5 5 3	号
学位授与の日付	2 0 2 3 年 9 月 2 5 日		
学位授与の要件	博士の論文提出者 (学位規則第 4 条第 2 項該当)		
学位論文の題目	Extension of the generalized lasso application in the spatial data analysis (空間データ解析における一般化 lasso 適用の拡張)		
論文審査委員	教授 坂本 亘	教授 飯塚 誠也	教授 石岡 文生
学位論文内容の要旨			
<p>Because of its flexibility in defining the penalty matrix and the predictor matrix, the generalized lasso becomes a potential method in spatial analysis. This study aims to explore and extend the application of generalized lasso in spatial data analysis.</p> <p>In the first part, we conducted simulation studies in the context of spatial clustering to evaluate some methods for selecting optimum tuning parameter of the generalized lasso. We defined the index of edges detection accuracy (IEDA), which indicates the accuracy of pooling coefficients. The approximate leave-one-out cross-validation (ALOCV) was found to be a recommended method compared to k-fold cross-validation, and GCV can be a substitution when ALOCV is not computable for relatively small tuning parameter or when the noise in the data has large variability.</p> <p>In the second part, we provided two applications of the generalized lasso for spatial clustering. First, we examined several methods of selecting the optimum tuning parameter for spatial clustering in the application to Chicago Crime Data. We obtained a suitable result of spatial clustering by using ALOCV, which selected the tuning parameter close to the one suggested by previous literature. Second, we extended the generalized lasso model for spatio-temporal clustering analysis, which can be separated into the two generalized lasso problems: trend filtering on the temporal scale, and fused lasso for spatial clustering. We applied the proposed method for detecting dynamic pattern of clusters on Covid-19 data in Japan. We obtained information on the clusters of prefectures, and how they are merged or dissolved.</p> <p>In the last part, we provided two applications of the generalized lasso for spatial modeling. First, we applied the generalized lasso to fit the spatially varying coefficient model and to cluster regional effects of socio-economics factors that affect the Covid-19 case in Java Island, Indonesia. We found that the poverty variable had no effect generally, while the HDI had different effect between regions. Second, we proposed a generalized lasso with two ℓ_1 penalties to fit a spatially varying coefficient model with both numerical and categorical predictor variables. The two types of penalties are used: for pooling regions and for pooling categories. Based on our simulation study and an application to the house sales price data in Java Island, our method could estimate coefficients well for pooling regions and categories.</p>			

論文審査結果の要旨

一般化lassoとは、回帰誤差の平方和にL1罰則項（回帰係数の絶対値の和）を加えて最小化することで変数選択などに利用できるlasso回帰を、係数の線形結合に対する罰則に拡張することにより、時空間上のクラスタリングや平滑化などに応用する方法である。本論文は空間データ解析における一般化lasso適用の拡張についての研究成果をまとめたものである。

本論文は三つの部分で構成される。第一に、一般化lassoを用いた回帰モデルの複雑さを制御する調整パラメータの選択方法を評価するためのシミュレーション研究を行った。空間構造をもつデータに対しては、k分割交差検証法の利用に問題があることから、近似1例消去交差検証法 (ALOCV) および一般化交差検証法 (GCV) の利用が推奨されることを示唆した。

第二に、空間クラスタリングにおける一般化lassoの二つの適用事例を与えた。シカゴ犯罪データへの適用では、ALOCVによる調整パラメータの選択で、先行研究で示唆された結果に類似する結果が得られた。日本のCovid-19感染者数データへの適用では、時空間クラスタリングへの拡張として、時間軸上のトレンドフィルタリングと空間上のfused lasso（隣接地点の係数の差に対する罰則を導入したもの）に分離できるモデルを提案し、都道府県間のクラスターの時間変化をヒートマップで視覚的に表現する結果が得られた。

第三に、空間変化係数モデルへの一般化lassoの適用に関する二つの適用事例を与えた。インドネシア・ジャワ島のCovid-19感染者数に影響する社会経済指標の地域効果のクラスタリングにより、人間開発指数 (HDI) の効果が地域間で異なることを示唆した。さらに、数値とカテゴリー両方の予測変数を含む空間変化係数モデルに対して、地域の併合のためとカテゴリーの併合のための2種の一般化L1罰則項を含む一般化lassoの適用を提案した。シミュレーション研究およびジャワ島の住宅売価データへの適用により、地域とカテゴリーの併合が適切に行われることを示した。

これらの研究成果は、査読付き雑誌に1編、国際会議論文集に2編掲載されており、環境統計学の分野への貢献は著しく、博士（環境学）の授与に値する研究であると判断した。