

2019年度 統計データ分析コンペティション

総務大臣賞（大学生・一般の部）

地方創生目標指標に関する変化要因ネットワークの 推定とそれに基づく地域間連携策の提案

張 瀚天、白鳥 友風

（筑波大学大学院システム情報工学研究科）

論文の概要

地方創生の設定目標について、目標指標に影響を与える変数の相互関係を明らかにするため、共分散行列を用いて要因ネットワーク図を作成し、lasso法による分析を行った。その結果、地域の稼ぐ力が地方創生に重要であることを指摘するとともに、地理的な制約にとらわれない地域間連携を提案している。

論文審査会コメント

グラフィカルモデリングを用いて、地方創生に関わる探索的關係性解析を多方面で実施し、データ分析能力が高く評価された。政策に関わる仮説実証型論文とは異なるが、近年のデータマイニングや機械学習に関わる方法論の政策科学分野への適用可能性を示した論文であり、得られた結論も探索的なものではあるが興味深いものもあった。

地方創生目標指標に関する変化要因ネットワークの推定と それに基づく地域間連携策の提案

張 瀚天・白鳥 友風

筑波大学大学院 システム情報工学研究科 社会工学専攻

1 はじめに

わが国における人口減少、少子高齢化に関する調査や対策の必要性への言及は枚挙に暇がない。特に地方において事態は深刻であり、例えば東京、埼玉、千葉、神奈川の人口の合計は2015年で日本の人口の約28.4%であるが、これは2010年の約27.8%から0.6%の増加、人数にして約51万3千人の増加となっている^[4]。国内における人口減少を勘案すると地方から大都市圏への人口流出が伺える。これらの課題に歯止めをかけるため、内閣府地方創生推進事務局は2015年に第1期「まち・ひと・しごと創生総合戦略」を策定した^[9]。この方針では、「地方創生」を達成するための基本目標として「1. 地方に仕事をつくり、安心して働けるようにする」、「2. 地方への新しいひとの流れをつくる」、「3. 若い世代の結婚・出産・子育ての希望をかなえる」、「4. 時代にあった地域をつくり、安心な暮らしを守るとともに、地域と地域を連携する」の4項目（以降、目標1～目標4と呼ぶ）を設定している。つまり、これらの基本目標を満足させることで「地域創生」が充足され、人口減少・少子高齢化の解決の糸口になるという方針が打ち出されている。さらに、これら目標の達成を定量的に評価するべく、まち・ひと・しごと創生総合戦略では、これらの基本目標に対して15件、それぞれの施策に関して116件のKPI (key performance indicator) が設定されている^[12]。

それらの基本方針達成に向けて様々な取り組みがなされており、「連携中核都市圏構想」はその1つである。連携中核都市圏構想は平成26年度から全国で展開されている制度であり、地方において中核都市が近隣の市町村と連携することによって圏域全体の経済を牽引し、人口減少や少子高齢化社会などの影響による縮小経済においても社会経済を維持するための施策である。これにより、2019年9月時点で連携中核都市圏として32圏域の宣言がなされている状況にある^[5]。一方で、先行研究により指摘があるように、連携中核都市圏に関する問題点として、圏内に含まれない地域があること、またそれに伴って圏外から圏内への人口流出を助長し、圏外の地域経済を悪化させる可能性があることが挙げられる^[13]。しかし、連携による経済効果も指摘されているように^[7]、現状のすべての地方に対して、より広範囲での地域連携によって地域経済が活性化されるような枠組みが必要であると考えられる。

また、基本目標はそれぞれ独立に達成しうるものではなく、様々な要素が互いに影響しあい、達成しうることに注意が必要である。例えば、目標1と目標2の関連で言えば、目標1に関連するその地域の産業の「付加価値額」は、目標2に関連する「人口流出率」と相関があることが報告されているが、同時に「人口流入率」と「所得水準」にも相関があり、人口と所得水準は循環して影響しあうとの指摘がある^[11]。これは目標1と目標2の要素が互いに循環して影響しあう例であり、このように互いの目標が影響しあい達成しうるものであると考えられる。

このように、基本目標の達成には、その地域の民間企業の雇用人数や、各地方自治体による教育の補助や財源の確保、地域間の連携など、複数の主体による直接的・間接的な要因が複雑に絡み合っており、基本目標を達成するということに対して各自治体は具体的にどのような対策を講じればよいのかが煩雑で不透明である状況となっている。そのため、各市町村において何を改善することが基本目標に含まれる要素の改善に繋がるのかを明らかにすることは地域創生の方針における根本的な課題といえる。これを踏まえ、本稿では各地域では「どのような指標が基本目標の達成につながるか」、「その指標を改善するためにはどのような対策が必要か」の2点を明らかにすることを目的とする。

2 分析方針と方法

2.1 分析方針

本稿では各市町村に関する包括的なデータセットを作成し、それによりどの要因が地方創生の達成に繋がりうるかを明らかにする。具体的には、まち・ひと・しごと創生総合戦略の4つの基本目標のKPIに深く関連する指標（以降では、目標指標と表記）を定義し、その指標の過去の変動要因を明らかにする。また、まち・ひと・しごと創生総合戦略では5年ごとに期が設定されており、2020年から第2期「まち・ひと・しごと創生総合戦略」が始動する^[10]。それを勘案し、5年における複数の指標の変化のデータを分析する。以降では具体的な分析案について述べる。

2.2 分析1：目標指標の増加率と人口規模の関連

前述の通り、各地方自治体では、都市部への人口が流出しており、日本全体の人口現状を鑑みると経済規模が徐々に縮小してしまっていると考えられる。一方で都市部では依然として人口が多く、地方と比べ経済規模も大きい。まち・ひと・しごと創生総合戦略に掲げられている通り、地方創生において人口が最も重要な変数の1つであると考えられる。分析1ではこのような人口と目標指標の変化量を相関分析をすることで、すでに人口規模が大きいところが目標指標を改善しやすいのか、もしくは違った傾向が存在するのかを明らかにする。

2.3 分析2：目標指標に影響する変数の特定

目標指標は他の様々な要因によって複雑に関係しあっていると考えられる。そこで分析2では、目標指標と他の様々な変数との関係を、ガウシアングラフィカルモデル(GGM)を用いたネットワーク分析によって推定する。GGMは、多変量正規分布に従う p 次元確率変数ベクトル $\mathbf{X} = (X_1, \dots, X_p)$ が与えられたとき、変数 X_i と変数 $X_j (j \neq i)$ の関係を条件付き分布を用いて推定し、グラフ化する手法である。そのモデルは、

$$X_j = - \sum_{k \neq j} \frac{\omega_{jk}}{\omega_{jj}} + \epsilon_j$$

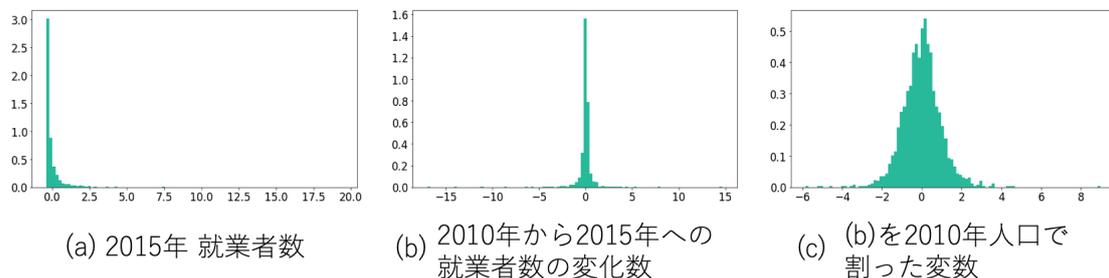
と表される。ここで、 $\omega_{i,j}$ は説明変数の共分散行列の逆行列の (i,j) 要素 $(i,j = 1, \dots, p)$ であり、 $\epsilon_j \sim N(0, 1/\omega_{jj})$ である。ここで、 $\omega_{jk} = 0$ と推定されれば、「 X_j と X_k は、他の説明変数が与えられた下で独立である」つまり「他の変数を考慮すると、 X_j と X_k は関係性がない」という条件付き独立と解釈することができる。実際の推定では計算誤差などの影響から ω_{jk} が0であると推定することは難しい、そのため関係性が弱いと考えられるものを0と推定するスパース推定法を合わせて適用するグラフィカルlasso (glasso)^[14]を用いる。

対象とするデータは複雑に影響しあっているため、多重共線性の問題が考えられる。GGMは線形回帰分析と同様、多重共線性と呼ばれる変数間に高い相関があることでパラメータ推定が不安定になるという問題が発生しうるが、glassoは頑健性を持つことが知られており^[1]、分析対象に適した分析手法であると考えられる。さらに分布を仮定することから、そのモデルの性能を尤度を用いて、AIC(赤池情報量基準)やBIC(ベイズ情報量基準)を計算することで客観的に評価することができる特徴がある。

加えて、推定された目標指標に関する変数間のネットワーク構造から、その構造においてどの変数が最も影響を及ぼしているかを、その構造においての媒介中心性^[16]を用いて評価することで目標指標に対する変数の重要度を評価する。媒介中心性はネットワークにおけるノード(変数)が他のノード間の最短距離のパスに含まれた回数で評価され、どれだけ他の変数に影響を与えているかを評価することができる。本稿ではglassoで推定された係数情報を活用する。係数が大きいということは関係が強いということを表すため、係数の逆数をそのノード間の距離として、各変数の媒介中心性を計測する。その際に、変数間の関係である係数が必ずしも正であるとは限らないため、正の係数からなるネットワークと負の係数からなるネットワークそれぞれを独立に評価し、各ノードに対して正の係数ネットワークでの媒介中心性から負の係数ネットワークでの媒介中心性を引くことで、そのノードの正の影響度合いを調査する。第4章にて後述するが、影響の解釈性を地方創生への貢献と一致させるため、失業者数などの減らすことを目標する変数に関しては符号を逆転させ解釈を一致させる。

図表 1: 使用した変数の基本統計量と出典, 及び使用した分析

項目名	データ年	分析番号	平均	分散	標準偏差	中央値	最頻値	出典
就業率	2010年	1	34109.774	8052.764	137.000	12151.000	170317.000	e-Stat(2010)
女性就業率	2010年	1	14608.194	36170.037	40.000	5299.500	687742.000	e-Stat(2010)
男性就業率	2010年	1	23400.623	19866.369	0.000	851.500	113969.000	e-Stat(2010)
雇用率(正社員の雇用・従業員[人])	2010年	1	17413.041	4366.465	77.000	6174.500	94829.000	e-Stat(2010)
2015年	1	25912.021	6095.056	14.000	1871.500	18176.000	e-Stat(2010)	
転出超過(転入・転出者数)	2010年	1	2916.079	8625.378	14.000	742.000	191862.000	e-Stat(2010)
2015年	1	1044.164	586.498	-1683.000	0.000	-51.000	563.000	収入した変数を元に計算
維持件数(転入)	2010年	1	400.889	1164.660	0.000	111.500	2205.000	e-Stat(2010)
2015年	1	613.190	1630.777	0.000	184.000	3093.000	e-Stat(2010)	
維持件数(転出)	2010年	1	12.401	23.658	0.000	0.000	791.000	e-Stat(2010)
2015年	1	337.254	8486.349	139.000	11813.500	1672013.000	SSDSE-2019A	
完全失業率(人)	2010年	1	14192.610	30627.510	48.000	5268.000	702607.000	SSDSE-2019A
2015年	1	1492.562	3916.208	0.000	508.500	67778.000	SSDSE-2019A	
雇用率(正社員の雇用・従業員[人])	2010年	1	17393.618	4362.369	85.000	6041.000	912607.000	SSDSE-2019A
2015年	1	2884.894	9481.111	18.000	838.000	183661.000	SSDSE-2019A	
転出超過(転入・転出者数)	2010年	1	2895.374	10008.196	12.000	697.500	198717.000	SSDSE-2019A
2015年	1	4.480	726.965	-3688.000	-62.000	11662.000	収入した変数を元に計算	
維持件数(転入)	2010年	1	363.736	1072.772	0.000	96.000	18634.000	SSDSE-2019A
2015年	1	215.629	1272.707	0.000	100.000	30027.000	SSDSE-2019A	
転入(人口)・変化数	2010年	1	13.864	29.634	0.000	6.000	621.000	e-Stat(2010)
2015年	1	-0.045	0.053	-0.270	-0.047	0.000	0.000	SSDSE-2019A
維持件数(転出)	2010年	2	-0.013	0.024	-0.154	-0.013	0.203	SSDSE-2019A
2010年-2015年	2	-0.002	0.012	-0.069	-0.002	0.043	SSDSE-2019A	
完全失業率(人)・変化数	2010年	2	-0.013	0.006	-0.067	-0.012	0.019	SSDSE-2019A
2010年-2015年	2	-0.005	0.006	-0.031	-0.005	0.017	SSDSE-2019A	
雇用率(正社員の雇用・従業員[人])・変化数	2010年	2	-0.001	0.006	-0.053	-0.001	0.136	SSDSE-2019A
2010年-2015年	2	-0.002	0.007	-0.052	-0.002	0.136	SSDSE-2019A	
転入者数(人)・変化数	2010年	2	-0.001	0.007	-0.070	-0.001	0.056	SSDSE-2019A
2010年-2015年	2	-0.001	0.001	-0.006	-0.001	0.011	SSDSE-2019A	
維持件数(転出)・変化数	2010年	2	0.001	0.001	0.011	0.001	0.017	SSDSE-2019A
2010年-2015年	2	0.001	0.001	0.004	0.001	0.017	SSDSE-2019A	
15歳未満人口(人)・変化数	2010年	2	0.000	0.000	-0.001	0.000	0.003	SSDSE-2019A
2010年-2015年	2	0.000	0.000	-0.004	-0.001	0.000	SSDSE-2019A	
15歳-64歳人口(人)・変化数	2010年	2	-0.057	0.032	-0.195	-0.050	0.147	SSDSE-2019A
2010年-2015年	2	0.027	0.022	-0.123	0.025	0.083	SSDSE-2019A	
65歳以上人口(人)・変化数	2010年	2	0.000	0.003	-0.012	0.000	0.041	SSDSE-2019A
2010年-2015年	2	0.000	0.000	-0.018	0.001	0.011	SSDSE-2019A	
高齢者人口(人)・変化数	2010年	2	0.007	0.004	-0.032	0.007	0.037	SSDSE-2019A
2010年-2015年	2	0.002	0.045	-0.746	0.000	0.558	SSDSE-2019A	
総労働力人口(人)・変化数	2010年	2	19.195	109.256	493.171	9.389	1993.197	e-Stat(2009,2014)
第1次産業従事者数(経済センサス・基礎調査結果) [人]・変化数	2009年-2014年	2	0.000	0.000	-0.006	0.000	0.003	SSDSE-2019A
第2次産業従事者数(経済センサス・基礎調査結果) [人]・変化数	2009年-2014年	2	-0.001	0.001	-0.014	-0.001	0.021	SSDSE-2019A
第3次産業従事者数(経済センサス・基礎調査結果) [人]・変化数	2009年-2014年	2	-0.003	0.004	-0.035	-0.003	0.024	SSDSE-2019A
先立金(歳入) [百万円]・変化数	2010年	2	1.492	7.989	-45.213	0.750	216.798	SSDSE-2019A
2010年-2015年	2	0.000	0.002	-0.033	0.000	0.035	SSDSE-2019A	
経常収支比率(市町村別) [%]・変化数	2010年	2	-0.001	0.002	-0.025	0.000	0.015	SSDSE-2019A
2010年-2015年	2	48.869	316.268	-604.073	26.765	9295.962	SSDSE-2019A	
児童労働者数(市町村別) [千人]・変化数	2010年	2	0.993	21.847	-307.656	0.516	456.352	SSDSE-2019A
2010年-2015年	2	62.864	291.937	-3704.400	24.827	5010.060	SSDSE-2019A	
地方労働力(市町村別) [千人]・変化数	2010年	2	25.424	67.609	-239.726	19.396	1717.203	SSDSE-2019A
2010年-2015年	2	12.927	100.785	-1710.629	1.413	1911.813	SSDSE-2019A	
生涯賃金(市町村別) [千円]・変化数	2010年	2	0.775	69.977	-506.686	0.335	1667.379	SSDSE-2019A
2010年-2015年	2	5.038	39.270	-142.412	0.000	152.863	SSDSE-2019A	
教育費(市町村別) [千円]・変化数	2010年	2	0.000	0.000	-0.001	0.000	0.000	SSDSE-2019A
2010年-2015年	2	0.000	0.000	-0.001	0.000	0.000	SSDSE-2019A	
小中学校数(市町村別) [校]・変化数	2010年	2	0.000	0.000	-0.002	0.000	0.000	SSDSE-2019A
2010年-2015年	2	0.000	0.000	-0.001	0.000	0.000	SSDSE-2019A	
高等学校数(市町村別) [校]・変化数	2010年	2	0.001	0.006	-0.011	0.000	0.000	SSDSE-2019A
2010年-2015年	2	0.000	0.000	-0.002	0.000	0.006	SSDSE-2019A	
公共施設数(市町村別) [校]・変化数	2010年	2	0.000	0.000	0.000	0.000	0.000	SSDSE-2019A
2010年-2015年	2	-0.650	0.074	-0.454	-0.033	0.860	e-Stat(2010,2015)	
小売店数(市町村別) [店]・変化数	2010年	2	0.001	0.001	-0.019	-0.001	0.008	SSDSE-2019A
2010年-2015年	2	0.000	0.001	-0.006	0.000	0.001	SSDSE-2019A	
飲食店数(経済センサス・基礎調査結果) [店舗数]・変化数	2010年	2	0.000	0.000	0.000	0.000	0.001	SSDSE-2019A
2010年-2015年	2	0.000	0.000	-0.001	0.000	0.000	e-Stat(2010,2015)	
大型小売店数(経済センサス・基礎調査結果) [店舗数]・変化数	2010年	2	0.000	0.000	-0.001	0.000	0.000	e-Stat(2010,2015)
2010年-2015年	2	0.000	0.000	-0.001	0.000	0.001	e-Stat(2010,2015)	
一般労働者数(市町村別) [千人]・変化数	2010年	2	0.000	0.000	-0.002	0.000	0.001	e-Stat(2010,2015)
2010年-2015年	2	0.000	0.000	-0.002	0.000	0.001	e-Stat(2010,2015)	
一般労働者数(市町村別) [千人]・変化数	2010年	2	3.954	43.073	-306.263	0.989	543.990	e-Stat(2010,2015)
2010年-2015年	2	4.515	219.218	-916.201	1.941	3040.985	e-Stat(2010,2015)	
投資的消費(市町村別) [百万円]・変化数	2010年	3	947966.795	4924840.593	123.000	18554.000	10971539.000	e-Stat(2009,2014)
2010年	3	3.298	6.703	0.000	1.959	61.077	RESAS(2010)	
従業員特化係数(産業) [国]	2016年	3	5.732	23.309	0.000	0.000	419.500	RESAS(2016)
従業員特化係数(製造業) [国]	2016年	3	3.915	13.993	0.000	0.000	206.938	RESAS(2016)
従業員特化係数(卸売業・小売業) [国]	2016年	3	4.432	0.993	0.000	1.226	8.622	RESAS(2016)
従業員特化係数(建設業) [国]	2016年	3	1.337	0.790	0.000	1.231	4.483	RESAS(2016)
従業員特化係数(情報通信業) [国]	2016年	3	1.227	4.890	0.000	0.242	96.960	RESAS(2016)
従業員特化係数(運輸業) [国]	2016年	3	0.174	0.461	0.000	0.053	6.658	RESAS(2016)
従業員特化係数(医療業) [国]	2016年	3	0.902	0.651	0.000	0.716	6.111	RESAS(2016)
従業員特化係数(教育業) [国]	2016年	3	0.895	0.228	0.189	0.912	1.972	RESAS(2016)
従業員特化係数(宿泊業) [国]	2016年	3	0.524	0.370	0.000	0.478	5.062	RESAS(2016)
従業員特化係数(不動産業) [国]	2016年	3	0.557	0.402	0.000	0.490	2.637	RESAS(2016)
従業員特化係数(情報・通信・娯楽サービス業) [国]	2016年	3	0.548	0.724	0.000	0.440	17.949	RESAS(2016)
従業員特化係数(電気・ガス・熱供給・水道業) [国]	2016年	3	0.996	0.704	0.016	0.866	6.793	RESAS(2016)
従業員特化係数(金融業) [国]	2016年	3	1.027	0.594	0.000	0.973	8.880	RESAS(2016)
従業員特化係数(学術・技術・サービス業) [国]	2016年	3	0.612	0.953	0.000	0.446	8.790	RESAS(2016)
従業員特化係数(医療・福祉) [国]	2016年	3	1.113	0.484	0.000	1.066	3.378	RESAS(2016)
従業員特化係数(教育サービス業) [国]	2016年	3	2.152	2.103	0.071	1.589	38.484	RESAS(2016)
従業員特化係数(サービス業) [国]	2016年	3	0.670	0.334	0.000	0.617	4.504	RESAS(2016)



図表 2: 分析 2 における前処理例

2.4 分析 3: 相乗効果が期待できる地域連携パターンの推定

分析 2 によって特定された目標指標に対して重要である変数に注目し、その変数値をもとにクラスタリングを行う。具体的には、分析対象データの性質を活かし、変数を離散化し、潜在クラス分析 [17] により、共通のクラスターに分類する。潜在クラス分析は、離散変数の尤度を EM アルゴリズムによって最大化することで、類似した変数の傾向をもつサンプルをクラスタリングするアルゴリズムである。尤度を用いた分析であるため、glasso 同様に AIC や BIC などでクラスタ数などの調整すべきパラメータを客観的に評価し決定することができる点で強みをもつ。さらにクラスタリングの結果はそのクラスターへの所属確率で表されるため、クラスタリングの結果が各サンプルに対して 1 つのラベルで返されるハードクラスタリング手法とは異なり、各サンプルに対して所属するクラスターを複数持つことができ、柔軟な地方連携を提案することができる。

本稿では、glasso と潜在クラス分析において、あらかじめ決めておかなければならないパラメータであるハイパーパラメータについてともに AIC を用いて客観的に評価を行う。glasso のハイパーパラメータであるスパース性に関わる正則化項の強度を調整するパラメータは 0 から 0.5 まで 0.001 間隔で探索を行い、潜在クラス分析のハイパーパラメータであるクラスター数は 2 から 50 までの間で探索を行いそれぞれ AIC が最小であったパラメータを選択する。

3 データ概要

本稿で使用したデータの基本統計量、出典、及びどの分析で用いたかを図表 1 に示す。本稿では、市町村に対して広範囲な変数を収録している SSDSE-2019A データを元に分析対象を決定したが、まち・ひと・しごと創生総合戦略に合わせ 5 年の変化量に着目するため、SSDSE-2019A データに存在する項目でも e-Stat から取得した変数も存在する。SSDSE, e-Stat, RESAS などの異なる出典からのデータ統合の際には市町村に対してのユニークな id である地域コードを元に統合し、欠損を除いた 1722 市町村に対して分析を行う。ただし、宮城県富谷市、福岡県那珂川市に関しては人口規模の増加により町から市へ変わったことから地域コードが合致しない箇所があるが、それらは同一の市町村とみなし分析を行う。

分析の際に 4 つの基本目標に対する KPI を元にそれらに深く関連する 10 項目を目標指標として設定する。具体的には目標 1 に関して、就業者数、女性就業者数、完全失業者数、正規雇用者数の 4 つの変数を、目標 2 に関して、転入者数、転出者数、転出超過数の 3 つの変数を、目標 3 として婚姻件数、出生数、保育所等数の 3 つの変数を設定する。目標 4 に関しては、KPI として設定されている項目が「立地適正化計画を作成する市町村数」など市町村にまたがる項目が多く単一市町村における分析と異なる点、収集できるデータの限界から本稿においては目標指標を設定しない。

分析 2 に関して、本稿で用いるガウシアングラフィカルモデルは、使用する変数に多変量正規分布を仮定している。しかし、使用データは図表 2(a), (b) に示す通り、元データ、2010 年と 2015 年の差の両方とも、正規分布裾野を厚くしたよう分布（ベキ分布）をしている。そのため、分析 1 の結果を踏まえ、より正規分布に近づけるための処理として、2010 年と 2015 年の差を、2010 年の各地域の総人口で割ることで一人当たりの値として扱う。図表 2 に目標指標の 1 つである就業者数に対して前処理の例を示す。使用する変数の中で、歳入の変化数、歳出の変化数、投資的経費の変化数は互いに 0.9 以上の相関係数があるため、ほぼ同一の変数を表していると考え、歳出の変化数、投資的経費を除いた。同様に総人口の変化数と 15~64 歳人口の変化数も 0.85 以上の相関係数であったため、総人口の変化数を除いた。また非水洗化人口は変化が負の方向にしか存在しなかったため、正規性の仮定から除いた。推定されるグラフの構造には影響しないが、推定された係数の符号の解釈と地方創生に対する影響を合致させるため、完全失業、転出者、転出超過、死亡者、経常収支比率、実質公債費比率の変化数の変数に関しては、-1 をかけることで符号を逆転させた。

分析 3 に関して、売上金額、従業員特化係数をクラスタリングを行う際に離散化を行う。これは、特化係数が特定の産業に対しての全国平均に対して、その地域のその産業の特化比率を表しており 1 を超えていることが大きな意味を持つと考えられること、さらに通常連続値でのクラスタリングを行う際にユークリッド距離などの距離を計測するが、特化係数の分布もまた 1 以下の値が極端に大きいベキ分布に従っているため距離的に近いサンプルが極端に多くなってしまい正しくクラスタリングをすることができないことを考慮し、これを緩和する目的で離散化を行う。具体的には規模を考慮するため、売上金額では 0.1 から 1 まで 0.1 刻みの分位点の範囲で含まれるかどうかにより離散化し、特化係数に関してはその値が 1, 3, 5 を超えているかどうかにより離散化する。

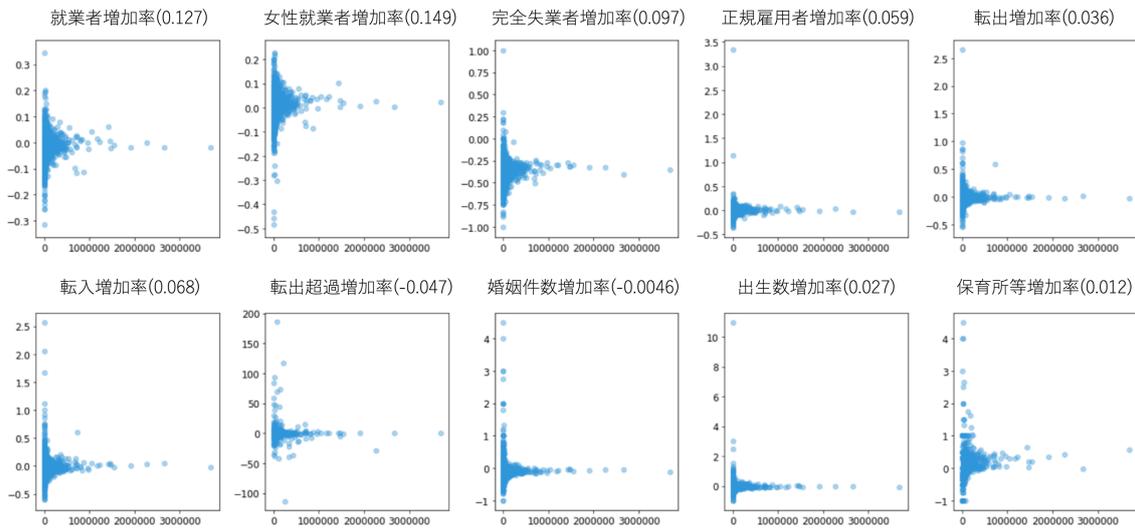
4 結果と提案

4.1 分析 1 の結果と解釈：目標指標に対する人口規模の影響

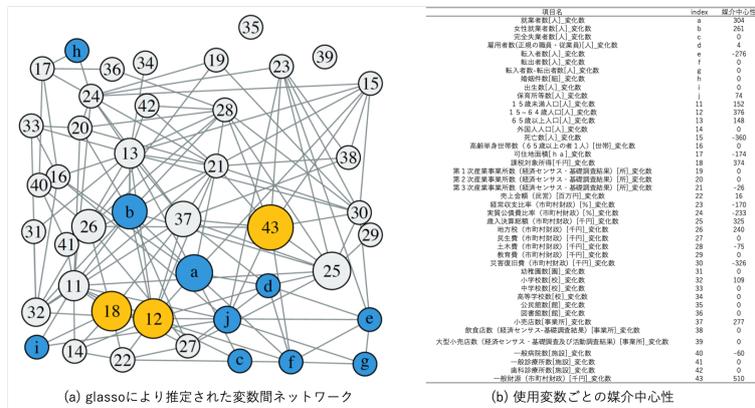
図表 3 は地域ごとに 2010 年総人口と各種目標指標の増加率を散布図としてプロットしたものである。いずれも相関係数は小さな値をとっており、また図からも比例関係が見られないことから、人口規模は目標指標の増加率に対して影響しないということが確認できる。このことから現在人口を集められている地方であっても、そうでなく過疎化が進行している地方であっても、過去 5 年間における目標指標の増加率が関係しないと考えられることが明らかになった。そのため、分析 2 ではこれらの全市町村に共通する目標指標の変動の要因を明らかにする。

4.2 分析 2 の結果と解釈：目標指標に影響する変数の特定

スパース推定法を用いたガウシアングラフィカルモデルである glasso により得られたネットワークを図表 4(a) に示す。10 の目標指標に関しては、ノードのインデックスとして小文字のアルファベットを用いて色を変更した。前述



図表 3: 分析 1: 2010 年総人口と各目標指標との相関係数と散布図

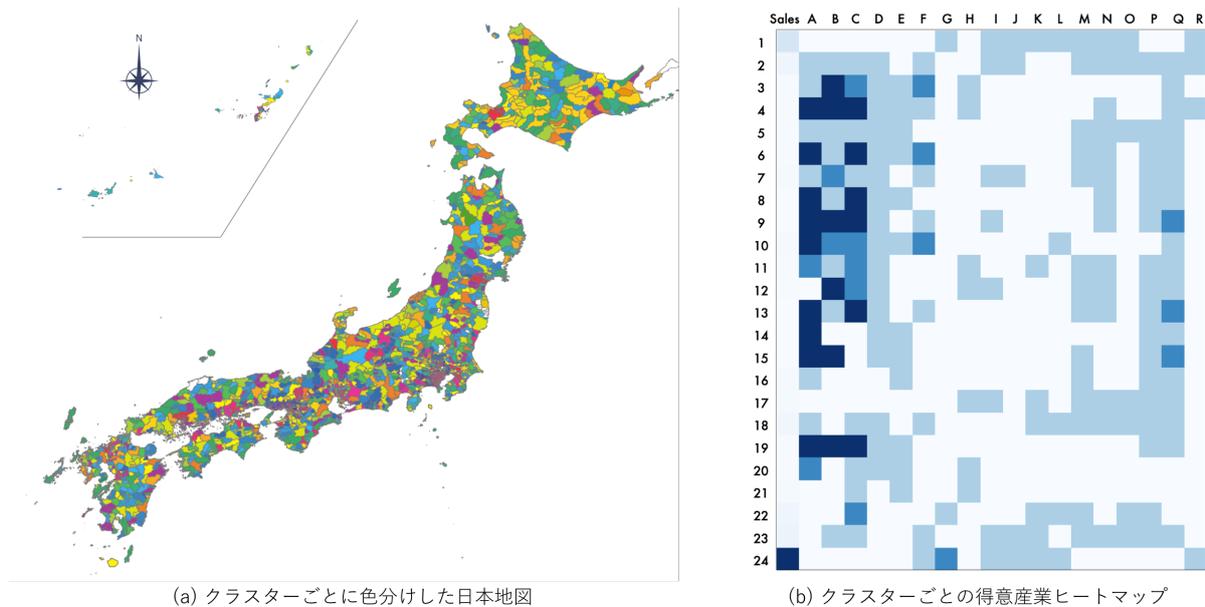


図表 4: (a)glasso により推定された変数間の関係ネットワーク: 小文字アルファベット a から j までは目標指標を表しており、ノードの大きさは媒介中心性の大きさを表している. (b) 推定されたネットワークにおける媒介中心性

の通りハイパーパラメータに関しては AIC が最小となる値に決定した. ネットワーク図において線で繋がっているノードは前述の通り、互いに影響を与え合う変数といえる. 推定されたネットワークから、各変数は互いに複雑に影響しあっていることがわかる、また目標指標である 10 の変数が全て含まれていることから、冒頭で述べた目標指標間の複雑な関連性が客観的に示されたといえる.

前述の通り glasso によって推定されたネットワークからそのグラフ全体への正の影響を評価した媒介中心性を計算する. 計算された各変数における媒介中心性は図表 4(b) に示されている. この媒介中心性の大きな変数に関しては、図表 4(a) のノードの大きさとして反映させており、その上位 3 つの変数はノードの色に反映させている.

最終的に正と負の係数ネットワークを考慮し、媒介中心性が最も高かった変数としては、順に市町村の一般財源、15~64 歳人口、課税対象所得、地方税、歳入決算額の変化数が挙げられた. これらは推定されたネットワークでその変化が他の変数におよぼす正の影響が強かった変数であると考えられる. ここで、前述の通り 15~64 歳人口は総人口と、歳入決算額は歳出決算額、投資的経費ととても強い相関をもつ変数であることに注意されたい. これらの変数は、多くが各自治体の経済規模の変化に関連する項目であり、中村良平 (2018)^[8] で指摘されているように地方自治体における「稼ぐ力」であると考えられる. これらの変数が最も大きな正の影響を与えるという結果は、地方創生において強調されている各地方自治体の「稼ぐ力」を表す変数であると考えられる. また推定されたネットワークは地方創生において、各自治体の「稼ぐ力」が経済の好循環を生む出発点である、という考えと合致する^[8]. 媒介中心性の負の値が大きかった変数としては、災害復旧費、実質公債費比率、経常収支比率などが挙げられる. 災害復旧費に関しては災害に見舞われた地方において高くなる傾向があるため、災害の影響から他の変数に対して負の影響を与えてしまっている



図表5: (a) AIC 基準で算出されたクラスターごとに色分けした日本地図, (b) 各クラスター (縦軸) の売上合計, 特化係数に関する平均値のヒートマップ: 特化係数に関しては視覚化のためそれぞれクラスター平均値が 1, 3, 5 を超えているか否かに応じて色を濃くしている。

ためだと考えられる。実質公債費比率, 経常収支比率に関しては, 分母と分子の与える影響がそれぞれ異なるのではないかと考えられる。例えば, 経常収支比率は分母である経常一般財源や, 分子である経常経費に充当した一般財源費などはそれぞれの変化によって他の変数に与える影響が異なるため, それぞれの影響が強くなり, より強かった負の影響が高くなったのではないかと考えられる。

分析2から, 複数存在した地方創生における目標指標は統一的な変数間の関係ネットワークにすべて含まれ, そしてそれらのネットワーク全体へ大きな正の影響を与える変数としては地方の「稼ぐ力」である経済規模の変化である変数が挙げられることが明らかになった。これらの分析から地方の「稼ぐ力」を向上させる案を考える。

具体的には, 本稿では目標4にもある「地方連携」を考える。理由は4つ存在する。1つめは, 地域によって得意とする産業や環境が異なるが, 同様の状況にある自治体と連携することでそれらの状況に適した案を創出できると考えるからである。2つめは, 各々の地方では今後増々の人口減少の傾向が考えられるため, 地方間で連携し経済規模を担保することが必要であるからである。推定されたネットワークにおいて15~64歳人口の変化量が2番目に正の影響を与える変数となった。これは総人口と相関が強い変数であるが, 今後人口の増加は増々見込めなくなっていく。そうした制約下で恒常的に他の一般財源などの増加を達成する, つまり「稼ぐ」には, 連携中核都市圏構想にあるように連携し, 経済規模を大きく育てることが必要不可欠であると考えられるからである。3つめは, 現在存在する連携中核都市圏構想では広域連携が難しい地域が多数存在する^[6]ことである。前述の通り連携中核都市圏構想では近隣の地域と連携するため, 連携が難しい地域が出てくる。しかし, 地方間の連携においては, 同様の問題を抱えている, 類似産業について強みをもっているなど地理的な制約のみにとどまらなると考えられる。4つめの理由として海外に向けての「稼ぐ力」を向上させる目的からである。日本の総人口は縮小傾向にあるため, グローバル市場に目を向けることは重要になってくる。これに関し, 先行研究において特定の産業において従業員数の規模と輸出量に相関があることが指摘されており, このように規模を獲得するためにも連携が必要であると考えられる^[15]。以上を踏まえ, 分析3では, 市町村データを横断的にクラスタリングを行う。

4.3 分析3の結果と解釈: 相乗効果が期待できる地域連携パターンの推定

クラスタリングを行うに当たり RESAS により取得できる特化係数をクラスタリングの変数として用いる。特化係数は大分類において18の各産業についてのその地域の強みや注力度合いを表した値であり, 従業者数, 付加価値額, 労働生産性の3つに対しての全国的な構成比に対しての比率で算出される。1を超えていればその産業に注力してい

各自治体に共通する変化要因のネットワークを明らかにした。推定されたネットワークでは各目標指標をすべて含んでおり、地方創生において各変数が複雑に影響しあっていること、そしてその構造が明らかになった。推定されたネットワークに対して各自治体の「稼ぐ力」に関連する変数群がとて大きな正の影響を持つ変数であることが明らかになり、現在のまち・ひと・しごと創生においての考えに対して分析による客観的な根拠を与えた。それらの結果を踏まえ、既存の地理的な制約にとらわれない地域連携に役立てるためクラスタリングを行った。

本稿では以上のように、地方創生に本質的に大きな影響をもつ要因を客観的な分析によって明らかにした上、その要因を改善するための案を示した点で、第2期の「まち・ひと・しごと創生総合戦略」の取り組みに大きな示唆を与える。

しかし、分析には大きく分けて2つの限界が存在した。1つは得られる収集できるデータの限界である。市町村単位での要因を明らかにするため、市町村単位のデータを取得する必要があるが、不足とする変数が存在した可能性がある。たとえば、5年でなく、1年ごとの変化や、4つの目標に即した施策が行われていたかなどの新たな変数が加わることでより詳細で精緻な分析ができたと考えられる。また、分析3での特化係数の欠損なども収集できるデータの限界であるといえる。2つめとして分析の将来のデータに対する妥当性である。本稿では用途、データに応じて分析手法を変えて適用した。そしてAICなどを用いることで客観的に妥当性を評価できる結果を導くことができた。ただし今回使用したデータは2016年以前のデータが多く地方創生に対して取り組みがまだ多くはなかった時期のデータである。そのため、各自治体によって様々な介入が行われた後の第2期に対してこの分析結果がどれほど一般性をもつのかは検証が必要である。ただし、分析自体はデータを変えることで新たな結果を導くことができるため、最新の国勢調査などの結果を用いることで分析をアップデートすることができると思う。

参考文献

- [1] 井手剛. (2010). 潜在的グラフ構造からの異常検知. Technical Report of the 1st Workshop on Latent Dynamics.
- [2] 経済産業省 (2016). 地域経済産業政策の現状と今後の在り方について. 地域経済産業グループ.
- [3] 厚生労働省 (2013). 失業率と他の経済指標の関係. 労働市場分析レポート., 第11号.
- [4] 総務省 HP. 人口推計. 総務省 HP (2019年9月8日時点).
- [5] 総務省 HP. 連携中枢都市圏構想. 総務省 HP (2019年9月8日時点).
- [6] 総務省 (2018). 広域連携が困難な市町村における補完のあり方に関する研究会報告書.
- [7] 中村良平 (2015). 地域創生に求められる地域経済構造分析 (特集 明日の地方創生を考える). 『土地総合研究』 23(3), 72-85.
- [8] 中村良平 (2018). 第4回 まちの稼ぐ力 (地方創生の本質).
<https://www.rieti.go.jp/jp/papers/contribution/hyogo-jichi/04.html>
- [9] 内閣府地方創生推進事務局 (2015). まち・ひと・しごと創生総合戦略 2015 改訂版.
- [10] 内閣府地方創生推進事務局 (2019). まち・ひと・しごと創生基本方針 2019 改訂版.
- [11] 三菱 UFJ リサーチ&コンサルティング (2016). 人口減少が地域経済に与える影響～商業、製造業が衰退、サービス業、農業が活性化のカギに～. シンクタンク・レポート.
- [12] まち・ひと・しごと創生本部 (2019). まち・ひと・しごと創生総合戦略の KPI の検証について (修正版)
- [13] 森川洋. (2016). 連携中枢都市圏構想の問題点について再度考える. 自治総研, 通巻, (457).
- [14] Friedman, J., Hastie, T., Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3), 432-441.
- [15] Wagner, J. (2007). Exports and productivity: A survey of the evidence from firm - level data. *World Economy*, 30(1), 60-82.
- [16] Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2), 163-177.
- [17] Linzer, D. A., Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of statistical software*, 42(10), 1-29.