

市区町村間の健康格差に関する統計分析

Statistical Analysis for Health Difference Among Municipalities

鳴原亮介*
Ryosuke SHIGIHARA

*地域計画学研究室 (指導教員: 奥村誠 教授)

It has become increasingly important for municipalities to choose their policies and concentrate their expenditure. We have to analyze complex relationships between many factors and needs of inhabitants. But there is no method which can analyze such relationships among too many factors. This paper propose a method to estimate true effects of each factor free from multicollinearity. The method is applied to municipal data concerning on health environments. This paper confirmed that there are no multicollinearities.

Key Words : *Multicollinearity, Principal Component Analysis, Weighted Regression, Health Difference*

1. はじめに

1.1 背景

公共政策に対する住民の需要が複雑化するなか、自治体は政策を選択し、財源を集中させることが必要になってきており、どのような要因をコントロールすれば住民の需要への対応を効果的に行うことができるのかを定量的に評価する必要がある。本研究では、ある政策目標に関連して、複数の指標が観測された多次元のデータセットから、特定の指標が他の指標に与える因果関係を推定する方法の開発を試みた。その際、通常回帰分析を行うと多重共線性が発生し(図1)、分析結果の信頼性を低下させてしまう。

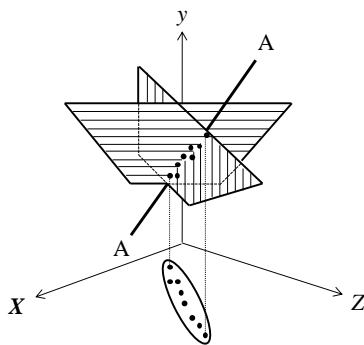


図1 多重共線性

1.2 既存研究

多重共線性の問題を解消し、説明変数が目的変数に与える真の影響を推定することができる手法として、Rosenbaum&Rubin¹⁾の「傾向スコアマッチング」、Rubin²⁾の「傾向スコアによる重み付け推定法」などが近年開発されている。しかし、これら既存の分析手法には、(a)傾向スコアは現象に決定的な影響を持つと考えられる1つの要因に着目したうえで各個体について算出される値であり、複数の要因が同時に影響を持つことが明らかである場合や、あらかじめどのような要因が大きな影響力を

持つかが明らかでない場合は適用することが難しい、(b)「自治体の健康さ」のように、それに影響を与えている要因はある程度判明しているものの、どれか1つに着目することはせずにこれらの要因を全て考慮したいという場合に適用することができない、という2つの問題点が挙げられる。

1.3 目的

本研究では、前述の2つの問題点を解消する新たな手法を開発することを目的とする。問題点(a)に対しては、関心のある複数の説明変数について算出した複数の傾向スコアを同時に考慮し重みとすることで、複数の説明変数が目的変数に与える真の影響を同時に推定することが可能となる「傾向スコアによる多重重み付け推定法」を、問題点(b)に対しては、特定の説明変数に着目することなく、説明変数を主成分に集約し、間接的に複数の説明変数に関連する重みを算出する「主成分得点による(多重)重み付け推定法」を提案する。そして、「健康な生活を送ることができる自治体」を目指した自治体が何をどのくらい整備すれば良いのかという観点から、全国の1927市区町村を対象とし、提案した「主成分得点による(多重)重み付け推定法」を用いて分析を行う。既存の手法と本研究で提案する手法との関係を表1に示す。

表1 既存の手法と提案する手法との関係

	分析単位	説明変数の数	要因	
傾向スコアマッチング	個体	一つ	明瞭	既存手法
傾向スコアによる重み付け推定法	群全体	一つ	明瞭	
傾向スコアによる多重重み付け推定量	群全体	複数(少数)	明瞭	提案手法
主成分得点による(多重)重み付け推定量	群全体	複数(多数)	不明瞭	

2. 単一要因に関する多重共線性の解消

いま各個体について、目的変数 y_i 、着目する説明変数 Z_i 、 Z_i 以外の説明変数 (以下、共変量と呼ぶ) X_i が観測されている。 Z_i が y_i に与える真の影響を推定するために、 Z_i について X_i で説明する傾向スコア e_i を算出し、 e_i の値

から個体 i について、 Z_i に関して特別な条件を与えられた個体群 (以下, 処置群), 特別な条件を与えられなかった個体群 (以下, 対照群) に割り当てを行う. 各個体が処置群に割り当てられる場合に $z_i > 0$, 対照群に割り当てられる場合に $z_i < 0$ とする Z_i に関する割り当て変数 z_i を導入する. 傾向スコア e_i は個体 i が処置群に割り当てられる確率として定義され, 式(1)のように表現される.

$$e_i \equiv p(z_i > 0 | x_i) \quad (1)$$

e_i を用いた 2 群でのマッチングにより, X と Z の相関が消え, 多重共線性が解消される. さらに, 各個体を 1 個の群に割り当てることを可能にした一般化傾向スコアの値を用いた 1 群でのマッチングも提案されている. その後の理論的研究を通じて指摘されたマッチングの欠点のうちいくつかを解決するために Rubin により「傾向スコアによる重み付け推定法」が開発された. 傾向スコア e_i の逆数を重みとして与えると, 処置群, 対照群全体での y の期待値 $E(y_+)$, $E(y_-)$ はそれぞれ,

$$E(y_+) = \sum_{i=1}^n \frac{\delta_i y_i}{e_i} / \sum_{i=1}^n \frac{\delta_i}{e_i} \quad (2)$$

$$E(y_-) = \sum_{i=1}^n \frac{(1-\delta_i)y_i}{1-e_i} / \sum_{i=1}^n \frac{1-\delta_i}{1-e_i} \quad (3)$$

と推定される. ただし, δ_i は $z_i > 0$ のとき 1, $z_i < 0$ のとき 0 となるダミー変数である. この操作においては, 分布が少ないところに存在する個体の重み (個数) を拡大している. そのため, X_i からみた処置群の個体の期待存在数が一様となり, XZ 平面上の個体の分布が一様となるため, X と Z の相関が消え, すなわち多重共線性を解消させることができる (図 2).

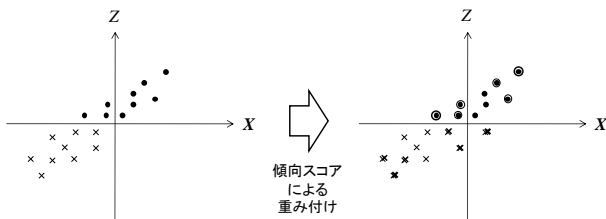


図2 傾向スコアによる重み付け推定法

しかし, 「1 群でのマッチング」や「傾向スコアによる重み付け推定法」には 1.2 節で述べた 2 つの問題がある.

3. 提案手法

3.1 傾向スコアによる多重重み付け推定法

問題点(a)に対して, 式(2), 式(3)で導入された重み (傾向スコア) の逆数の代わりに, 複数の説明変数について算出された傾向スコアを掛け合わせたものの逆数で置き換える. これにより, 複数の説明変数に関する多重共線性の解消を目指す. 処置群, 対照群全体での y の期待値 $E(y_+)$, $E(y_-)$ はそれぞれ,

$$E(y_+) = \sum_{i=1}^n \prod_j \frac{\delta_{ij} y_i}{e_{ij}} / \sum_{i=1}^n \prod_j \frac{\delta_{ij}}{e_{ij}} \quad (4)$$

$$E(y_-) = \sum_{i=1}^n \prod_j \frac{(1-\delta_{ij})y_i}{1-e_{ij}} / \sum_{i=1}^n \prod_j \frac{1-\delta_{ij}}{1-e_{ij}} \quad (5)$$

と推定される. ただし, δ_{ij} は j 番目の説明変数について $z_{ij} > 0$ のとき 1, $z_{ij} < 0$ のとき 0 となるダミー変数である.

3.2 主成分得点による (多重) 重み付け推定法

主成分分析では, 互いに相関を持つ多数の指標 $x = (x_1, x_2, \dots, x_p)$ を互いに無相関な少数の主成分 $Z = (Z_1, Z_2, \dots, Z_m) (m \leq p)$ に集約させることができる. また, 個体 i について主成分得点が算出される.

$$Z_i = a_{1i}x_1 + a_{2i}x_2 + \dots + a_{pi}x_p \quad (6)$$

問題点(b)に対して, 主成分得点から算出される重みを用いることで特定の変数に着目することなく間接的に多数の変数について同時に着目することができるという考えから, 「主成分得点による (多重) 重み付け推定法」を提案する. この操作によって X と主成分に集約された複数の説明変数との多重共線性を解消し, 主成分に集約された複数の説明変数が y_i に与える真の影響を推定することが可能になる.

重みの計算はまず, 個体 i について算出された主成分得点 Z_i を標準化した標準化主成分得点の度数分布図を作成する. 各階級の度数はその階級に属する個体の数を表しており, その階級値に対応する標準化主成分得点を持つ個体の出現確率を表している. 各個体に対して第 1 主成分得点の出現確率の逆数 $1/w_1$ を重みとして与えることで, 第 1 主成分の軸上でみた全ての階級における個体の期待存在数が一様になるように調整される. このとき, y の期待値は式(7)のようになる.

$$E(y) = \sum_{i=1}^n \frac{z_i y_i}{w_1} / \sum_{i=1}^n \frac{z_i}{w_1} \quad (7)$$

図 3 のように, この段階では XZ 平面上で見ると第 1 主成分の軸上に個体が偏在していることには変わりはなく, XZ 間の相関に起因する多重共線性は解消されていない.

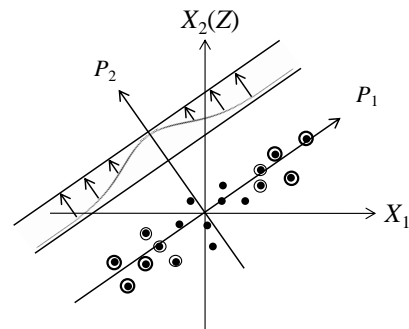


図3 主成分得点による重み付け (第 1 主成分得点まで)

そこで, 3.1 節で説明した多重重み付け推定法を用い, さらに第 2 主成分得点の出現確率 w_2 も同時に考慮し, w_1 と掛け合わせたものの逆数 $1/(w_1 w_2)$ を重みとして与える. このとき, y の期待値は以下のようになる.

$$E(y) = \sum_{i=1}^n \frac{z_i y_i}{w_1 w_2} / \sum_{i=1}^n \frac{z_i}{w_1 w_2} \quad (8)$$

この操作では図 4 のように、第 1, 第 2 主成分の平面上において個体の期待存在数が一樣になるように調整され、その結果 XZ 平面上で見た場合についても期待存在数が一樣となるため、結果として X と Z との多重共線性が解消される。このように、第 l 主成分について多重共線性を解消させるためには、第 $(l+1)$ 主成分得点の出現確率の逆数まで重みとして考慮する必要がある。

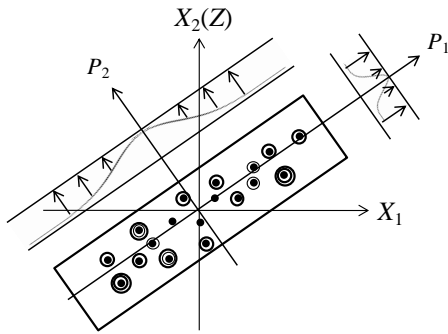


図 4 主成分得点による重み付け(第 2 主成分得点まで)

多重共線性の検出方法として、各説明変数に対して算出される分散拡大要因(VIF)という指標が一般的に利用されている。VIF は当該説明変数に対して単回帰をとった場合の標準誤差と重回帰をとった場合の標準誤差の比として定義され、 $\sigma_{\hat{\beta}_i}^2/\sigma_{\hat{\beta}_i}^2 = (x^T x)^{-1}$ のように表される。しかしこの指標は、通常回帰分析に対応しているが重みを考慮した回帰分析には対応していない。そこで本研究では、重み付き回帰分析に対応した VIF を次のように算出し、分析における多重共線性の発生の様子を調べた。

$$\sigma_{\hat{\beta}_i}^2/\sigma_{\hat{\beta}_i}^2 = (X^T W X)^{-1} X^T W W X^T (X^T W X)^{-1} \quad (9)$$

ただし、 W は重みの対角行列である。

4. 健康格差の分析

4.1 分析対象と使用する健康指標

「健康な生活を送ることができる自治体」を目指す自治体にとってどのような要因をコントロールすることが効果的であるかという問題に対し、「主成分得点による(多重)重み付け推定法」を適用した分析を行った。分析対象は全国の 1927 市区町村である。「健康さ」は死亡率 (= 死亡数/人口総数) を用いて表現し、目的変数とする。また「健康さ」を決定する説明変数は既存研究を参考に、自治体の政策で対応可能な範囲に含まれるであろう 58 の指標を使用した。

4.2 分析結果・考察

主成分分析で得られた結果について述べる(紙面の都合上、結果の掲載は割愛する)。主成分の寄与率は第 1 主成分が 32.43%、第 2 主成分が 7.92%、第 3 主成分が 5.79%、第 4 主成分が 5.39% となり、第 4 主成分までで全体の情報量の 51.54% を説明していることがわかった。各主成分の固有ベクトルから、第 1 主成分は「自治体の所

得水準」を表す主成分と判断した。同様に第 2 主成分は「生産地・消費地という自治体の性質」を表す主成分、第 3 主成分は「自治体における医療環境の充実・不足の程度」を表す主成分、第 4 主成分は「自治体の気候」を表す主成分である判断した。

次に主成分得点から算出した重みを考慮し回帰分析を行った結果を述べる。医療環境の整備が死亡率に与える真の影響を推定するため、第 1~第 4 主成分から算出される重みを同時に考慮した結果を表 2 に示す。また、比較のため重みを考慮しない通常回帰分析を行った結果を表 3 に示す。表 2、表 3 を比較すると、重みを考慮することで t 値の絶対値が大きく、有意な結果を示す説明変数が多くなった。また自由度調整済み決定係数の値に改善が見られた。VIF の値はいずれの説明変数についても値が小さくなる傾向が見られ、多重共線性が解消されている。医療に関わりのある説明変数を見ると、表 3 では出生人口あたり乳児死亡率、人口あたり薬剤師数は正に、緊急医療施設アクセス時間、人口あたり基本健康診査受診率は負に有意な要因となっている。表 2 では多重共線性のために、出生人口あたり乳児死亡率、人口あたり薬剤師数が死亡率に与える正の影響は表 3 と比べて過小に、緊急医療施設アクセス時間、人口あたり基本健康診査受診率が死亡率に与える負の影響は表 3 と比べて過大に推定されていることがわかった。また、表 2 では有意であった人口あたり医師数は表 3 では有意ではなくなった。

表 3 の t 値を見ると、死亡数に与える影響が医療に関わる説明変数よりも強いものが見つかった。人口あたり外国人数は死亡数に負の影響を持っていると推定された。これは、日本に住む外国人の年齢が相対的に低いためではないかと考えられる。財政力指数、人口あたり小売店数、人口あたり商業年間販売額はいずれも死亡率に正の影響を与えると推定された。これは、忙しさに伴い死亡率が高くなるためであると考えられる。人口あたり役員数、人口あたり飲食店数は自治体の豊かさを表す説明変数であると考えられ、死亡率に負の影響を与えていると推定された。総面積あたり道路実延長、人口あたり乗用自動車保有台数のような交通の便を表す説明変数は、死亡率に正の影響を与えていると推定された。全説明変数のうち最大の t 値が算出された人口あたり刑法犯認知件数は、死亡率に正の影響があると推定された。

「健康な生活を送ることができる自治体」を目指す自治体はこれらの結果から、死亡率を効果的に低下させる政策を判断することができる。

5. おわりに

本研究では複数の指標が観測された多次元のデータから、特定の指標が他の指標に与える偏りのない真の影響を推定する手法を提案し、全国の市区町村に対して提案手法による分析を行った。そして、「健康な自治体」とな

表2 回帰分析結果 (第1~第4主成分を重みとして考慮した場合)

説明変数	係数	標準誤差	t値	有意性	VIF
切片	-1.41E+01	1.30E+00	-10.873	***	-
人口・世帯					
log(人口あたり高齢者数)	7.16E-01	6.57E-02	10.897	***	5.91E-11
log(人口あたり外国人数)	-4.35E-02	4.88E-03	-8.929	***	3.26E-13
log(平均年齢)	9.45E-01	1.85E-01	5.112	**	4.68E-10
log(人口密度)	-4.62E-02	1.77E-02	-2.612	**	4.29E-12
log(人口あたり転入率)	-1.02E-02	2.37E-02	-0.431		7.66E-12
log(人口あたり在住率)	4.01E-01	6.31E-02	6.354	***	5.44E-11
行政基礎					
log(財政力指数)	2.12E-01	1.94E-02	10.957	***	5.13E-12
教育					
log(最終学歴率大学大学院)	-7.49E-03	2.30E-02	-0.326		7.21E-12
log(最終学歴率未就学)	-4.54E-03	4.79E-03	-0.948		3.14E-13
log(教員一人あたり小学校児童数)	-9.98E-02	2.01E-02	-4.963	***	5.54E-12
労働					
log(人口あたり平均週間就業時間)	-9.28E-02	4.22E-02	-2.197	*	2.44E-11
log(人口あたり完全失業率)	-1.74E-02	1.49E-02	-1.169		3.05E-12
log(人口あたり役員率)	-1.81E-01	1.84E-02	-9.868	***	4.63E-12
log(人口あたり第1次産業就業者率)	2.21E-02	5.90E-03	3.753	**	4.76E-13
log(人口あたり第2次産業就業者率)	1.23E-01	1.93E-02	6.374	***	5.10E-12
log(人口あたり第3次産業就業者率)	1.73E-01	4.44E-02	3.888	***	2.70E-11
文化・スポーツ					
log(人口あたり公民館数)	6.14E-03	1.79E-03	3.424	***	4.40E-14
log(人口あたり図書館数)	-1.32E-02	2.35E-03	-5.593	***	7.57E-14
居住					
log(人口あたり小売店数)	1.85E-01	1.80E-02	10.241	***	4.44E-12
log(人口あたり飲食店数)	-8.45E-02	9.10E-03	-9.287	***	1.13E-12
log(人口あたり大型小売店数)	-1.01E-02	3.56E-03	-2.845	**	1.74E-13
log(人口あたり百貨店数)	3.04E-03	2.16E-03	1.406		6.41E-14
log(人口あたり非水洗化人口率)	2.75E-02	4.23E-03	6.507	***	2.45E-13
log(人口あたり投票率)	3.48E-02	5.76E-02	0.603		4.55E-11
log(主要道路総延長あたり舗装率)	2.12E-01	1.10E-01	1.918		1.67E-10
log(総面積あたり道路実延長)	-7.08E-02	9.57E-03	-7.406	***	1.25E-12
log(可住地面積あたり新設着工住宅件数)	3.27E-02	6.13E-03	5.329	***	5.14E-13
log(可住地面積あたり金融機関店舗数)	1.19E-03	7.09E-03	0.28		6.88E-13
log(人口あたり預貯金残高)	-4.07E-02	6.40E-03	-6.362	***	5.60E-13
log(人口あたり乗用自動車保有台数)	-2.04E-01	2.56E-02	-7.991	***	8.96E-12
log(人口あたりテレビ契約台数)	-1.59E-04	2.45E-02	-0.006		8.24E-12
log(可住地面積あたり郵便局数)	5.30E-02	1.80E-02	2.951	**	4.42E-12
log(可住地面積あたり都市公園数)	-1.00E-02	3.58E-03	-2.809	**	1.75E-13
経済基礎					
log(人口あたり課税対象所得)	1.16E-01	1.82E-02	6.398	***	4.52E-12
log(人口あたり商業年間商品販売額)	5.52E-02	7.08E-03	7.788	***	6.87E-13
log(人口あたり製造品出荷額)	-3.57E-03	3.59E-03	-0.994		1.77E-13
log(人口あたり農業算出額)	-3.67E-03	1.59E-03	-2.302	*	3.47E-14
自然環境					
log(総面積あたり可住地面積率)	1.51E-02	2.11E-02	0.719		6.07E-12
log(年平均気温)	-2.18E-01	4.60E-02	-4.73	***	2.89E-11
log(年平均相対湿度)	6.95E-01	1.62E-01	4.285	***	3.61E-10
log(年間日照時間)	-4.11E-01	8.56E-02	-4.795	***	1.00E-10
log(年間降水量)	8.00E-02	2.98E-02	2.684	**	1.22E-11
log(年平均風速)	-2.82E-02	2.14E-02	-1.316		6.28E-12
log(最深積雪)	-3.83E-03	2.59E-03	-1.478		9.20E-14
福祉・社会保障					
log(人口あたり国民健康保険被保険者率)	3.44E-01	5.16E-02	6.656	***	3.65E-11
log(可住地面積あたり保育所数)	2.92E-02	3.88E-03	7.533	***	2.06E-13
log(可住地面積あたり老人ホーム数)	9.65E-03	1.96E-03	4.926	***	5.25E-14
安全					
log(人口あたり建物火災出火件数)	-6.01E-03	2.43E-03	-2.478	*	8.05E-14
log(人口あたり交通事故発生件数)	-5.28E-02	8.45E-03	-6.249	***	9.79E-13
log(人口あたり刑法犯認知件数)	1.23E-01	7.69E-03	15.941	***	8.09E-13
健康・医療					
log(緊急医療施設アクセス時間)	-1.74E-02	5.47E-03	-3.172	**	4.10E-13
log(人口あたり実質医療費)	2.55E-02	5.21E-02	0.488		3.72E-11
log(人口あたり基本健康診査受診率)	-4.43E-02	1.20E-02	-3.708	***	1.96E-12
log(人口あたり一般病院数)	-2.52E-03	2.92E-03	-0.864		1.17E-13
log(人口あたり一般診療所数)	3.38E-04	1.21E-02	0.028		2.00E-12
log(出生人口あたり乳児死亡率)	4.41E-02	9.36E-03	4.71	***	1.20E-12
log(人口あたり薬剤師数)	1.53E-02	2.45E-03	6.241	***	8.24E-14
log(人口あたり医師数)	-4.37E-03	2.70E-03	-1.62		9.95E-14
自由度調整済み決定係数			0.9657		
標本数			1927		

注: .は10%, *は5%, **は1%, ***は0.1%水準で統計的に有意であることを示す

表3 回帰分析結果 (重みを考慮しない場合)

説明変数	係数	標準誤差	t値	有意性	VIF
切片	-9.61E+00	9.78E-01	-9.827	***	-
人口・世帯					
log(人口あたり高齢者数)	9.68E-01	6.08E-02	15.909	***	1.50E-01
log(人口あたり外国人数)	9.48E-03	5.19E-03	1.828	*	1.09E-03
log(平均年齢)	-1.20E-01	1.52E-01	-0.788		9.35E-01
log(人口密度)	-4.42E-03	1.41E-02	-0.314		7.99E-03
log(人口あたり転入率)	2.86E-02	1.89E-02	1.514		1.44E-02
log(人口あたり在住率)	5.16E-01	8.63E-02	5.98	***	3.01E-01
行政基礎					
log(財政力指数)	2.11E-02	1.55E-02	1.362		9.73E-03
教育					
log(最終学歴率大学大学院)	-5.52E-02	2.03E-02	-2.721	**	1.67E-02
log(最終学歴率未就学)	8.97E-03	3.63E-03	2.474	*	5.32E-04
log(教員一人あたり小学校児童数)	-9.65E-04	2.09E-02	-0.046		1.77E-02
労働					
log(人口あたり平均週間就業時間)	-3.19E-02	2.36E-02	-1.351		2.26E-02
log(人口あたり完全失業率)	8.33E-02	1.47E-02	5.678	***	8.71E-03
log(人口あたり役員率)	-2.27E-02	1.65E-02	-1.375		1.10E-02
log(人口あたり第1次産業就業者率)	2.63E-02	9.25E-03	2.845	**	3.46E-03
log(人口あたり第2次産業就業者率)	-5.45E-02	2.36E-02	-2.31	*	2.25E-02
log(人口あたり第3次産業就業者率)	-2.10E-03	4.50E-02	-0.047		8.21E-02
文化・スポーツ					
log(人口あたり公民館数)	1.42E-03	1.63E-03	0.87		1.08E-04
log(人口あたり図書館数)	5.87E-04	1.54E-03	0.381		9.57E-05
居住					
log(人口あたり小売店数)	8.82E-02	2.16E-02	4.082	***	1.89E-02
log(人口あたり飲食店数)	-4.01E-02	9.78E-03	-4.095	***	3.87E-03
log(人口あたり大型小売店数)	-7.27E-04	1.81E-03	-0.401		1.33E-04
log(人口あたり百貨店数)	-1.05E-03	1.77E-03	-0.594		1.27E-04
log(人口あたり非水洗化人口率)	7.07E-03	4.03E-03	1.755	*	6.57E-04
log(人口あたり投票率)	-2.65E-02	4.29E-02	-0.618		7.46E-02
log(主要道路総延長あたり舗装率)	-1.38E-02	7.62E-02	-0.181		2.35E-01
log(総面積あたり道路実延長)	-1.56E-02	1.09E-02	-1.431		4.81E-03
log(可住地面積あたり新設着工住宅件数)	2.26E-03	5.55E-03	0.406		1.25E-03
log(可住地面積あたり金融機関店舗数)	2.67E-03	4.76E-03	0.56		9.18E-04
log(人口あたり預貯金残高)	-2.43E-02	7.35E-03	-3.312	***	2.18E-03
log(人口あたり乗用自動車保有台数)	-8.64E-02	3.25E-02	-2.659	**	4.27E-02
log(人口あたりテレビ契約台数)	2.24E-02	3.73E-02	0.6		5.62E-02
log(可住地面積あたり郵便局数)	1.01E-02	1.20E-02	0.842		5.87E-03
log(可住地面積あたり都市公園数)	-1.71E-03	1.66E-03	-1.026		1.12E-04
経済基礎					
log(人口あたり課税対象所得)	1.16E-01	3.71E-02	3.137	**	5.57E-02
log(人口あたり商業年間商品販売額)	2.91E-03	8.89E-03	0.328		3.20E-03
log(人口あたり製造品出荷額)	8.57E-03	3.27E-03	2.617	**	4.34E-04
log(人口あたり農業算出額)	-9.79E-03	3.54E-03	-2.766	**	5.07E-04
自然環境					
log(総面積あたり可住地面積率)	-1.34E-02	1.72E-02	-0.784		1.19E-02
log(年平均気温)	-7.48E-02	3.76E-02	-1.99	*	5.71E-02
log(年平均相対湿度)	-2.65E-02	1.13E-01	-0.234		5.19E-01
log(年間日照時間)	-1.91E-01	6.88E-02	-2.77	**	1.92E-01
log(年間降水量)	-8.54E-03	2.21E-02	-0.387		1.97E-02
log(年平均風速)	2.30E-02	1.50E-02	1.533		9.09E-03
log(最深積雪)	-4.36E-03	1.78E-03	-2.443	*	1.29E-04
福祉・社会保障					
log(人口あたり国民健康保険被保険者率)	3.00E-03	4.65E-02	0.065		8.74E-02
log(可住地面積あたり保育所数)	8.14E-03	3.74E-03	2.174	*	5.67E-04
log(可住地面積あたり老人ホーム数)	9.20E-03	1.60E-03	5.75	***	1.03E-04
安全					
log(人口あたり建物火災出火件数)	3.92E-03	2.01E-03	1.951	*	1.63E-04
log(人口あたり交通事故発生件数)	1.67E-02	7.66E-03	2.184	*	2.38E-03
log(人口あたり刑法犯認知件数)	5.93E-02	7.99E-03	7.414	***	2.58E-03
健康・医療					
log(緊急医療施設アクセス時間)	-4.06E-03	5.43E-03	-0.748		1.19E-03
log(人口あたり実質医療費)	4.18E-02	3.53E-02	1.185		5.03E-02
log(人口あたり基本健康診査受診率)	-8.96E-03	9.29E-03	-0.965		3.49E-03
log(人口あたり一般病院数)	-9.94E-04	1.82E-03	-0.546		1.34E-04
log(人口あたり一般診療所数)	-2.56E-04	8.76E-03	-0.029		3.11E-03
log(出生人口あたり乳児死亡率)	1.63E-02	7.22E-03	2.263	*	2.11E-03
log(人口あたり薬剤師数)	1.56E-03	2.37E-03	0.657		2.28E-04
log(人口あたり医師数)	-1.72E-02	3.93E-03	-4.383	***	6.24E-04
自由度調整済み決定係数			0.8024		
標本数			1927		

注: .は10%, *は5%, **は1%, ***は0.1%水準で統計的に有意であることを示す

るためにはどのような指標をコントロールすれば効率的に達成することができるのかについて定量的な評価を行った。

今後の課題として、第1に分析に使用する個体数の問題が挙げられる。提案手法では、個体が存在しないほど外れた階級に対しては重みを計算することができないため、モデルへの当てはまりに問題が生じる可能性がある。そこで、使用する個体数を増やすといった工夫が求められる。第2に、重みとして考慮する主成分の数を決める明快な指針を開発することができていないという問題が挙げられる。多重共線性が発生していると回帰分析の設計行列の逆行列を計算できないという観点から、例えば

主成分の数を増やすたびに設計行列の固有値を確認するといった方法が考えられるが、詳細については今後の課題としたい。

参考文献

- 1) P. R. Rosenbaum, D. B. Rubin : The Central Role of the Propensity Score in Observational Studies for Casual Effects, *Biometrika*, Vol.70, pp41-55, 1983.
- 2) D. B. Rubin : The Use of Propensity Scores in Applied Bayesian Inference, *BAYSIAN STATISTICS 2*, pp 463-472, 1985.

(2011年2月7日提出)