

卒業論文

メロディーから ギターのコード譜を作成する リアルタイムシステム

Data Fusion through Web-GIS Visualization
Using Open Data for Evidence-Based Policy Making

富山県立大学 工学部 情報システム工学科

2020043 山内拓海

指導教員 Antonio Oliveria Nzinga Rene 教授

提出年月: 2023年2月

目次

図一覧	ii
表一覧	iii
記号一覧	iv
第1章 はじめに	1
§ 1.1 本研究の背景	1
§ 1.2 本研究の目的	1
§ 1.3 本論文の概要	2
第2章 従来の研究	3
§ 2.1 メロディーを表現する既存のシステム	3
§ 2.2 ギターに関する専門知識	5
§ 2.3 リアルタイムでのタブ譜作成の必要性	5
第3章 音楽分析のアプローチと印象の変化	7
§ 3.1 2つの楽譜の数値化による比較	7
§ 3.2 メロディーから音階のリアルタイム抽出	7
§ 3.3 コードの違いによる印象の変化	8
第4章 提案手法	9
§ 4.1 歌 (楽譜 A 付き) から楽譜作成 (B)	9
§ 4.2 楽譜からタブ譜作成のプログラム	12
§ 4.3 楽曲からタブ譜作成システムの開発	15
第5章 数値実験並びに考察	19
§ 5.1 数値実験の概要	19
§ 5.2 実験結果と考察	21
第6章 おわりに	24
謝辞	25
参考文献	26

図一覧

2.1	音楽情報検索の例	4
2.2	GIS を用いることの3つの利点	4
4.1	csv のフォーマット	11
4.2	システムのフロントページ	11
4.3	因果探索の結果とアウトプットの例	12
4.4	DEA 部分のフロー	13
4.5	DEA のアウトプットの例	15
4.6	EBPM-GIS の実装	17
4.7	提案システム	17
5.1	射水市の推計人口 [22]	20
5.2	年少・老年人口割合推移 [22]	20
5.3	入力データと出力データ	22

表一覧

4.1	地理情報を持たない数値データ	10
4.2	地理情報を持つ数値データ	10
4.3	位置データ	10
4.4	各都道府県の DMU 内訳	13
5.1	「年少人口 [割合]」に対する Direct-LiNGAM の結果	22
5.2	DEA の入力・出力一覧	22
5.3	元データと改善案の比較	23
5.4	参照の市区町村とウェイト	23

記号一覧

以下に本論文において用いられる用語と記号の対応表を示す.

用語	記号
LiNGAM における i 番目の観測変数	x_i
LiNGAM における j 番目の観測変数から i 番目の観測変数へのパス係数	b_{ij}
LiNGAM における i 番目の観測変数に対する誤差 (非観測変数)	e_i
主問題における各入力に対する重み	v^T
主問題における各出力に対する重み	u^T
主問題における対象 DMU の評価値	z
CCR モデルにおける DMU _o の入力	x_o
CCR モデルにおける DMU _o の出力	y_o
CCR モデルにおける DMU の入力	X
CCR モデルにおける DMU の出力	Y
双対問題における対象 DMU の評価値	w
入力指向モデルにおける対象 DMU の評価値	θ
入力指向モデルにおける各 DMU に対する重み	λ
出力指向モデルにおける対象 DMU の評価値	η
出力指向モデルにおける各 DMU に対する重み	μ
入力指向モデルにおける対象 DMU の i 番目の入力に対する改善案	$\hat{x}_i$
入力指向モデルにおける参照集合内の k 番目の DMU の i 番目の入力	x_{ik}
入力指向モデルにおける参照集合内の k 番目の DMU に対する重み	λ
出力指向モデルにおける対象 DMU の j 番目の出力に対する改善案	$\hat{y}_j$
出力指向モデルにおける参照集合内の k 番目の DMU の j 番目の出力	y_j
出力指向モデルにおける参照集合内の k 番目の DMU に対する重み	μ
提案手法における d 番目の市区町村の i 番目の入力	x_{id}
提案手法における d 番目の市区町村の i 番目の出力	y_{id}
提案手法における d 番目の市区町村に対する重み	λ_d
<i>robust Z-score</i> における正規化後の値	ι
<i>robust Z-score</i> を用いて正規化するデータ集合内の値	x
<i>robust Z-score</i> を用いて正規化するデータ集合	X
<i>robust Z-score</i> を用いて正規化するデータ集合の中央値	$median(x)$
<i>robust Z-score</i> を用いて正規化するデータ集合の正規四分位範囲	$NIQR$
0~1 変換の結果の値	ι'
0~1 変換を行うデータ集合内の値の最大値	$max \iota $

はじめに

§ 1.1 本研究の背景

音楽の創造的な表現は、人類が持つ豊かな文化の一翼を担い、感情やアイデンティティの表現において重要な役割を果たしています。音楽家やアーティストは、耳コピや即興演奏のスキルを駆使して、独自の音楽を生み出し、聴衆との共感を生み出しています。しかし、これらのスキルを磨く過程は熟練を要し、時には限定的な情報やリソースに依存することがあります。

従来の音楽理論や楽器の技術向上だけでなく、デジタル技術の進化により、新たな音楽制作の手法やサポートツールが登場しています。しかし、依然として楽曲の細部まで把握し、自らのアイデアを具現化するための効果的な手段が求められています。例えば、複雑なコード進行や微細なフレーズを正確に理解することは、アーティストが自分の音楽に深みを与える上で重要です。

現代の音楽制作において、アーティストやミュージシャンは多岐にわたる音響情報を取り込み、それを元に独自の音楽表現を構築しています。しかし、そのプロセスは手作業や経験に頼る部分が依然として存在し、特に耳コピや即興演奏においては、正確かつ迅速な情報処理が求められています。この課題に対処するために、デジタル技術を駆使した新しい手法やツールの開発が期待されています。

§ 1.2 本研究の目的

本研究の目的は、音楽制作における耳コピと即興演奏のプロセスを効果的に支援するためのデジタル技術の活用に焦点を当てます。特に、高速フーリエ変換（FFT）などのデジタル信号処理技術を活用し、楽曲中の周波数成分を精密に解析する手法を研究します。これにより、アーティストは楽曲の構造やフレーズを詳細に理解し、より表現力豊かな音楽を生み出すことが可能となります。

具体的には、以下の点に焦点を当てます。

耳コピの支援: 音楽の聴き取りや再現において、楽曲中のコードやメロディを高精度かつ迅速に抽出する手法の開発。

即興演奏の支援: リアルタイムでの音楽情報の解析と提供により、演奏中においても正確な音楽的情報を得ることができるシステムの構築。

これにより、アーティストは感性豊かな音楽制作がより容易になり、個々の表現力を最

大限に引き出すことができるでしょう。本研究により、デジタル技術の進歩が音楽制作に新たな次元をもたらす手助けとなることを期待しています。

§ 1.3 本論文の概要

本論文は次のように構成される。

- 第1章 本研究の背景と目的について説明する。音楽分野における即興演奏や耳コピの重要性について述べる。目的は背景で挙げた課題に対して、リアルタイムでの音高推定を用いて演奏支援システムの提案について述べる。
- 第2章 音楽分野における従来研究についてまとめ、リアルタイムでの tab 譜のについて述べる。また、ギターについての名称や専門知識についての説明を行う。
- 第3章 本研究の提案手法に用いる音楽理論の一部について解説する。また、提案手法のシステム部分に用いる音階の抽出方法について解説する。
- 第4章 本研究の提案手法について述べる。
- 第5章 実際にシステムを使用してもらい評価を行った。ギターを演奏してもらい有効性の検証を行った。
- 第6章 本論文における前章までの内容をまとめつつ、本研究で実現できたことと今後の展望について述べる。

従来の研究

§ 2.1 メロディーを表現する既存のシステム

音楽の生成と合成

音楽のメロディー分析は、多岐にわたる音楽関連の分野で重要な役割を果たしており、その重要性はますます高まっています。この分野では、音楽理論、楽曲制作、音楽情報検索、音楽教育などが相互に補完し合い、進化し続けています。

音楽理論の文脈では、メロディーは楽曲の中で聴衆に最も直接訴えかける要素の一つです。メロディーの分析は、異なる時代やジャンルの楽曲におけるメロディーの構造や進行を理解する手段となっています。音楽理論家や研究者は、メロディーにおける音程やリズムのパターン、そしてハーモニーとの関係性を分析し、作曲家がどのように聴衆の感情に訴えかけるかを解明しています。楽曲制作の現場では、デジタル音楽ワークステーション (DAW) や楽譜ソフトウェアが広く活用されています。これらのソフトウェアは、メロディー分析を通じて作曲家やプロデューサーに対して革新的な手段を提供しています。メロディーの自動生成や編集、アレンジの支援など、様々な機能が導入され、アーティストたちはこれらのツールを駆使して創造的なプロセスを進めています。

音楽情報検索システムは、大規模な音楽データベース内で楽曲を見つける際にメロディー分析を駆使しています。これにより、ユーザーは特定のメロディーを持つ楽曲を容易に見つけたり、関連する楽曲を発見したりできます。アルゴリズムや機械学習の進歩により、より高度なメロディー認識が可能となり、ユーザーエクスペリエンスが向上しています。音楽教育分野では、メロディー分析が学習者が楽曲を理解し、演奏するための基盤となっています。教育用ソフトウェアや教材は、学習者が楽曲内のメロディーの構造や音楽理論を効果的に理解できるよう工夫されています。また、メロディーの分析を通じて、学習者が創造的な表現力を養う手段ともなっています。総じて、これらの分野でのメロディー分析は、音楽における知識の深化や技術の進歩に寄与しています。これにより、アーティストや研究者、教育者はより豊かな音楽体験を提供し、次世代の音楽愛好者やプロフェッショナルに向けて新たな可能性を切り拓いています。既存のシステムについて他にも紹介する。

音楽情報検索とデジタル音楽ワークステーション

音楽のメロディー分析は、音楽理論や楽曲制作、音楽情報検索など、音楽研究のさまざまな分野で重要な役割を果たしている。これに関連して、多くの既存のシステムが開発および研究されている。既存のシステムの一例として、音楽情報検索分野におけるシステムが



図 2.1: 音楽情報検索の例

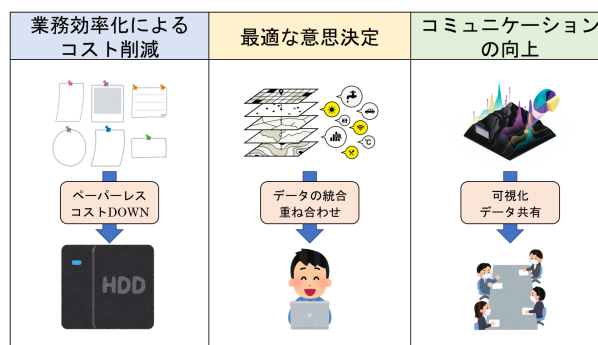


図 2.2: GIS を用いることの 3つの利点

挙げられる．これらのシステムは，楽曲のメロディーラインを自動的に抽出し，音楽データベース内で楽曲を検索したり，ジャンルを分類したりするのに使用されている．

さらに，音楽制作の分野では，デジタル音楽ワークステーションソフトウェアや楽譜ソフトウェアがメロディー分析のために使用されている．これらのシステムは楽曲制作プロセスにおいてメロディーの作成，編集，アレンジに必要な機能を提供し，ミュージシャンや作曲家に創造的なツールを提供している．また，音楽教育の分野でもメロディー分析が重要である [2]．

最後に，音楽制作プロセスにおいてメロディー分析は，アーティストやプロデューサーにとっても重要である．メロディーラインの解析は，楽曲の構造や楽器アレンジに関する洞察を提供し，音楽の創造的なプロセスを向上させる．メロディーを分析する既存のシステムは，音楽研究，制作，教育の各分野で多くの利点をもたらし，これらのシステムの進化は，音楽の理解や制作プロセスの向上に寄与しており，今後の音楽技術の発展に期待が寄せられている．

AIとメロディー生成

近年，人工知能（AI）の進化が音楽制作とメロディー生成の分野において注目を集めています．AIは機械学習やニューラルネットワークを活用し，既存のメロディーや楽曲から学習し，新しいメロディーを生成する能力を持っています．音楽制作において，AIはアーティストやプロデューサーに革新的な手段を提供しています．AIモデルは大量の楽曲データを分析し，異なるジャンルやスタイルに基づいてメロディーを生成することができます．これにより，創造的なプロセスが効率的になり，新しい音楽のアイデアが生み出される可能性が広がります．

一方で，AIのメロディー生成は音楽理論や感性の側面においても議論的となっています．アートの創造性や感情の表現において，AIが本当に人間と同様の深さを持つことができるのか，といった問いかけがなされています．これらの課題は，AIとメロディー生成の関係性において新たな研究やディスカッションを呼び起こしています．さらに，AIを活用した音楽情報検索システムも進化しています．AIは高度なメロディー認識技術を駆使して，ユーザーが特定のメロディーを持つ楽曲を見つけたり，関連する楽曲を推薦したりする際に役立っています．これにより，音楽のディスカバリー体験が向上し，ユーザーが新しい音楽に出会いやすくなっています．

総じて、AIとメロディー生成の結びつきは、音楽の創造的な領域や技術の進歩に新たな可能性をもたらしています。アートとテクノロジーが融合する中で、今後ますます興味深い展望が広がることでしょう。

§ 2.2 ギターに関する専門知識

ギターに関する専門知識において、弦とフレットについての基本的な理解は非常に重要です。ギターを演奏するためには、これらの要素について正確に理解しておく必要があります。

ギターは、その特有の音色と多彩な奏法で知られる楽器です。ギターの演奏を始める際に最初に覚えるべき重要な要素の一つが、弦とフレットの呼び方です。これらの用語を理解することは、ギター演奏の基礎知識として不可欠です。

まず、ギターの弦について説明しましょう。ギターは通常、6本の弦を持っています。これらの弦は、ギターのボディからネック（指板を含む部分）に伸び、ヘッド（ギターの先端）に取り付けられています。一番細い弦から順に1弦、2弦、3弦、4弦、5弦、6弦と呼ばれます。1弦は通常、高音域の音を担当し、6弦は低音域の音を奏でます。これらの弦は異なる太さと音程を持っており、演奏時にそれぞれ異なる音を出します。

次に、フレットについて説明しましょう。フレットは、ギターのネック（指板）に打ち込まれている金属のバーで、弦を押さえて異なる音を出すために使用します。フレットは、ヘッド（ギターの先端）に近いものから順に1フレット、2フレット、3フレット、と数えられます。この数え方により、どのフレットを押さえるかを簡単に指定できます。例えば、「1フレットを押さえる」という指示は、1番近いフレットを押さえることを意味します。

これらの基本的な用語を理解すると、ギターの演奏においてコードや音符を指示するための「TAB譜」も理解しやすくなります。TAB譜は、6本の弦の上にフレット数を書いた図で表現されます。例えば、TAB譜で「6弦3フレットと5弦5フレットを押さえる」という指示を見た場合、これは6番目の弦を3フレット目で押さえ、5番目の弦を5フレット目で押さえることを意味します。これにより、特定のコードやメロディーを演奏する際にどのフレットを押さえ、どの弦を鳴らすかが明確に示されます。

ギターはその多彩な奏法と音楽の表現力において魅力的な楽器であり、弦とフレットの理解はその才能を開花させるために欠かせない要素です。ギターを学び始めると、これらの基本的な概念が演奏技術の向上に貢献し、新たな音楽の世界を探求する手助けとなることでしょう。

§ 2.3 リアルタイムでのタブ譜作成の必要性

リアルタイムでのタブ譜作成は、音楽制作、演奏、学習、および共有において重要なツールである。このアプローチの必要性は多岐にわたり、ミュージシャンは演奏中に楽譜を作成し、アイデアを瞬時に記録でき、これにより、創造的なプロセスが向上し、楽曲制作が効率的に進行する。

また、学習者にとってもタブ譜は貴重なツールである。楽譜を見ながら演奏することで、

正確なフレットと弦の指示が得られ、スキルの向上が可能である。リアルタイムでのタブ譜は、音楽の学習を助け、演奏の正確性を高めると考えられる。リアルタイムでのタブ譜作成は音楽制作、演奏、学習、共有において重要なツールであり、ミュージシャンと音楽愛好者に多くの利点をもたらすと考えられる。

音楽分析のアプローチと印象の変化

§ 3.1 2つの楽譜の数値化による比較

因果探索とは、観測データを用いて、そのデータ群の因果グラフ（複数の観測データにおいて、それぞれの値がお互いに及ぼしあっている影響の度合いを構造的に示したもの）を導出するための教師なし学習のことである。

また、類似する手法として因果推論が挙げられるが、因果推論では因果関係の向きが既知である場合にその因果関係が本当に有意であるのかをデータから分析する手法であるのに対し、因果探索は因果関係が不明かつ因果関係の向きも不明であるデータ群に対して、それらの間に因果関係が成立するかを導く手法である。

例えば、「ある小売店 A でアイスクリームの安売りを行った際にアイスクリームの売り上げが向上した。また、同日の小売店 A の来客数は前日より 100 人多かった」というケースがあったとする。このとき、アイスクリームの安売りを行ったことが売り上げの向上につながったかどうかを調べるのが因果推論である。これに対して、アイスクリームが安かったから来客数が増加したのか、来客数が多かったためにアイスクリームの売り上げが向上したのかという因果の方向性も含めて分析を行うのが因果探索である。

このような特徴を持つため、因果探索は適用されるデータの分野に対しての制約が少なく、様々な分野のデータに適用することができる。それゆえ、因果探索を用いた応用研究も盛んにおこなわれており、疫学、経済学、神経科学、化学、医学をはじめとした幅広い分野のほか土木計画学 [13] の研究でも用いられている。

§ 3.2 メロディーから音階のリアルタイム抽出

魔王魂のピアノ 2-6 ラ、ガレージバンドのピアノのラ、人の声のの音源を用意し、FFT を用いたピーク検出により音高推定を行った。ピーク検出の様子と周波数、推定した音高を出力し、これを図 2 に示す。それぞれのデータについて考察を行った。魔王魂のピアノ 2-6 ラ、ガレージバンドのピアノのラの音源に対してはどちらも周波数が 879Hz を示し、音階では A5 と推定した。これは周波数一覧に示されているように妥当であると言える [5]。

また、人の声の音源では周波数がおよそ 214Hz を示し、音階では A3 と推定した。3 つの音源に対しても同じくラの音として推定できていることがわかる。同じラの音でも A3 と A5 と表示されているが、これは 2 オクターブの差があることを示していると考えられる。これらの判定をリアルタイムで演奏者に的確な情報を提供することを目的とする。

§ 3.3 コードの違いによる印象の変化

音楽において和音は、異なる高さの音が3つ以上組み合わせられたもので、その構成や構成音の間隔によって印象が形成されます。和音の感覚的な要素として、構成音間の半音の数や和音の安定性が挙げられます。

感覚的な要素として、構成音間の半音の数や和音の安定性が挙げられます。音楽理論では、これらの感覚を半音の数やファンクションに基づいて理論化しており、例えば半音メジャーコードやマイナーコードが「明るい-暗い」を表現し、ドミナントモーションが「安定-不安定」を形成します。

和音の感覚は一般的に、「明るい-暗い」や「安定-不安定」など共有されるものがあります。これらの感覚は音楽理論で、「明るい」を半音メジャーコードやマイナーコード、「安定」をドミナントモーションなどで表現しています。しかし、地域や文化によってはこれらの感覚が異なります。地中海沿岸では、特に短調の曲においてマイナー・メジャー・セブンスを主音とする感覚があるとされています。

地域や文化により、和音の印象が異なります。例えば、地中海沿岸では特有の感覚が報告されており、これらの差異を理解することで地域ごとの音楽文化や感性の違いが明らかになります。異なる文化や地域の感性に対する理解は、音楽の普遍性と地域差のバランスを考える上で重要です。

和音は楽曲において進行を導く要素として重要です。特定の和音の組み合わせやファンクションが、楽曲における安定性や不安定性を形成します。たとえば、ドミナントからトニックへの進行は楽曲における解決の瞬間を生み出し、聴衆に感動を与えることがあります。

和音の印象が聴衆に与える影響や異なる和音の組み合わせが感性にどう訴えかけるかについても考察します。明るい印象を与えるメジャーコードと暗い印象を持つマイナーコードの対比や、和音の高低の配置が聴衆に与える心理的な影響について深く掘り下げます。また、和音の進行が楽曲において期待感や緊張感をどのように生み出すかを具体的な楽曲例を挙げて解説します。

音楽の和音はその構造や構成によって印象が変わり、地域や文化による感性の違いを理解し、音楽理論を通じて和音の印象を解明することで、楽曲制作や音楽理解に新たな視点が得られることが示唆されます。和音の理解を通じて、音楽が人々の感性や文化とどのように対話しているのかを明らかにすることが、本研究の主要な貢献と言えるでしょう。

提案手法

§ 4.1 歌 (楽譜 A 付き) から楽譜作成 (B)

本研究の提案手法は、大別すると以下のような4つの工程からなる。

1. Web 上からあらかじめ収集し、サーバ上のデータベースに保存されたデータの中から利用者が任意のデータを選択する。
2. 選択されたデータとの間に因果関係が成立するデータを LiNGAM によってデータベース上から抜き出し、DEA の入力と出力に振り分ける。
3. 振り分けられたデータを用いて DEA を解き、評価値および、改善値を算出する。
4. DEA の結果を Web-GIS 上に可視化し、対象のデータと因果関係があり、地理的要素を持つデータについてはレイヤーを作成して重ね合わせが行えるようにする。

そのため、提案手法は以下のような4つのプログラムに分けることができる。

1. 対象のデータの受け取り
2. LiNGAM による因果分析およびデータの振り分け
3. DEA によるデータ分析
4. 結果の表示および重ね合わせ

本節では、ユーザが対象データを指定するにあたって前段階として必要なデータベースの作成方法とデータベースに保存されているデータの種類、対象の指定から LiNGAM による因果探索でのデータの選定、DEA の入力・出力に対するデータの振り分けまでの提案手法のしくみを解説する。

データベースに用いた統計データ

データベースに用いたデータの項目については表??に示す。データの属性としては位置データか数値データか、地理的情報を持つか否かで大きく3つに分けることができる。地理的情報を持たない数値データの例には人口、製造品出荷額などが挙げられ、数年に1回、全国統一でデータの収集が行われるものである。地理的情報を持つ数値データとしては、公園や医療機関の数が挙げられ、これらのデータに関してはどこにその施設が存在するかの位置データが紐づけられている。

本研究に用いるデータベースを作成するにあたって、政策の対象となる事柄について、その原因となるものを探すという目的を達成するために、ある程度広い分野の情報がそろっ

表 4.1: 地理情報を持たない数値データ

データ項目	単位	データ項目	単位
耕作放棄地率	%	経営耕地面積	1 畝 / 経営体
農業産出額	千万円	労働生産性	なし
企業数	社	従業員数	人
歳出決算額 [総務費]	%	農地平均取引価格	円 / m^2
歳出決算額 [民生費]	%	商業用地平均取引価格	円 / m^2
歳出決算額 [衛生費]	%	住宅用地平均取引価格	円 / m^2
歳出決算額 [農林水産業費]	%	林地平均取引価格	円 / m^2
歳出決算額 [商工費]	%	マンション等平均取引価格	円 / m^2
歳出決算額 [土木費]	%	1 人あたりの地方税	千円
歳出決算額 [警察費・消防費]	%	製造品出荷額	万円
歳出決算額 [教育費]	%	事業所数	事業所
歳出決算額 [公債費]	%	総人口	人
歳出決算額 [労務費]	%	老年人口	%
歳出決算額 [その他 (雑費)]	%	生産年齢	%
農業就業人口平均年齢	歳	年少人口	%
農業経営者平均年齢	歳	年間商品販売額	百万円
林作業請負収入	万円	海面漁獲物等販売額	万円
林産物販売金額	万円	付加価値額	万円
一人当たりの法人住民税	千円	1 人あたりの固定資産税	千円

表 4.2: 地理情報を持つ数値データ

データ項目	単位
施設数 [空港]	箇所
施設数 [工業団地]	箇所
施設数 [都市公園]	箇所
施設数 [道の駅]	箇所
施設数 [学校]	箇所

表 4.3: 位置データ

データ項目	単位
施設位置 [空港]	経度・緯度
施設位置 [工業団地]	経度・緯度
施設位置 [都市公園]	経度・緯度
施設位置 [道の駅]	経度・緯度
施設位置 [学校]	経度・緯度

ていること、日本全国に対応しており市区町村単位の粒度のデータであること、各地方や都道府県などでデータにおいて大きな欠損がないことなどの条件を満たす必要があった。

これらの条件から、地理的情報を持たないデータに関しては 2.1 節で挙げた RESAS の API を用いてデータの収集を行った。また、地理的情報を持つデータに関しては国土交通省の国土数値情報ダウンロードにおけるデータを収集し、それらに基づいて各市区町村ごとの施設数を自動的に数え上げることで施設数のデータとした。

全てのデータは csv 形式でサーバに保存し、各処理ごとに取り出して参照、編集、加工できるようにした。データベースとして用いた csv のフォーマットについて、各市区町村における人口および公園の数、位置のデータを例に挙げて図 4.1 に示す。RESAS によって収集したデータと施設の数を表すデータに関しては、1 列目に総務省から各市区町村に対して割り当てられている全国地方公共団体コードをキーとして格納し、2 列目に対象のデータを格納するという形をとった。施設の位置データについては

システムのフロントページおよび対象の選択

提案手法においてユーザサイドに提示されるフロントページを図 4.2 に示す。図 4.2 に示

人口		公園の数		公園の位置	
city_Code 総人口		city_Code	park_num	lon	lat
1100	1952356	1100	2661	143.3149	42.93257
1202	265979	1202	307	141.3449	42.99603
1203	121924	1203	92	141.9418	43.52811
1204	339605	1204	361	141.5346	43.06934
1205	88564	1205	112	144.2413	43.9901
⋮		⋮		⋮	
47381	3998	47381	0	127.8759	26.69031
47382	1843	47382	0	127.7192	26.217

図 4.1: csv のフォーマット

[illegible]

図 42: システムのフロントページ

した通り、システムのはじめにユーザには前述のデータベース内に格納されている全てのデータの項目名が提示される。また、データの項目名をクリックすることによってデータの中身を参照することができる。このページにおいてユーザは自身が対象としたい市区町村およびデータ項目を選択する。

データの前処理

表 4.1 にあるように本研究で分析に用いるデータは単位が異なり、値の大きさも広範囲にわたるため、分析を行う前にデータの正規化を行う。まず、本研究で用いるデータは正規分布に従わないため、そのようなデータに対しても適用することが可能な robust Z-score を用いる [20]。robust Z-score について以下に示す。

 $\langle \text{robust Z-score} \rangle$

$$\iota = \frac{x - median(x)}{NIQR} \quad (4.1)$$

robust Z-score では、式 (4.1) に示すように各データとデータ集合の中央値との差をとり、その値とデータ集合の正規四分位範囲との商を求めることによって値を正規化する。正規四分位範囲とは、四分位範囲と 1.3489 の商である。また、robust Z-score の結果では各値が 0 を中心に正規化されるため負の値となるという場合が発生する。しかし、CCR モデルによる分析では負の値を持つデータを扱うことが出来ないため、さらに値がすべて 0~1 の範囲に収まるように正規化を行う。用いる処理を以下の式に示す。

< 正規化 >

$$\iota' = \frac{\iota + \max |\iota|}{2 \max |\iota|} \quad (4.2)$$

因果探索によるデータの選定

本研究では、1.1 節で挙げた政策の対象における要因の複雑性という課題の解決を支援する手法として 3.1 節で挙げた LiNGAM による因果探索を提案する．具体的には、多量かつ幅広い分野のデータに対して LiNGAM による因果探索を適用することによって、政策の対象としたいターゲットと因果関係のあるデータのみを抜き出すという手法である．

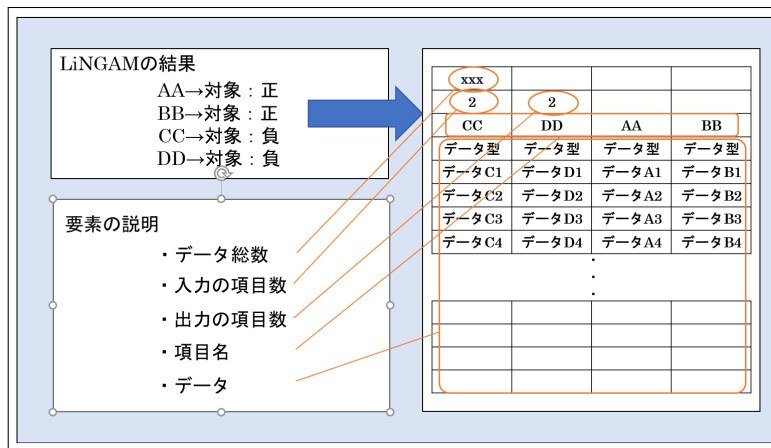


図 4.3: 因果探索の結果とアウトプットの例

因果探索における数学的なアルゴリズムに関しては 3.1 節で示した Direct-LiNGAM に従う。システムのアルゴリズムとしては、前述のフロントページにてユーザーが選択した政策の対象に関係するデータを Direct-LiNGAM を解くコードに送信することで、そのデータをターゲットとして因果探索を行う。

この際、因果探索に用いられるデータは表 4.1 および 4.2 に示したすべてのデータであり、これらすべてのデータを用いて一度に因果探索を行うことで政策の対象と潜在的に因果関係を持つデータのみを自動的に絞り込む。

また、これらの処理の結果を用いて 4.2 節にて後述する DEA を用いた分析を行うために、因果探索を解いた結果、対象のデータとの因果関係が示されたデータのみを結合した新たな csv ファイルを作成する。この際に DEA における入力・出力の振り分けも同時に行う。

本研究の提案手法では因果探索によって同定された因果グラフのうち、対象のデータに向く矢印を持つデータのみを分析の対象として扱うこととする。これは、本研究における因果探索の役割が DEA の入力・出力となるデータを絞り込むことにあるためである。因果探索の結果のうち、矢印の始点にあるデータが増減すると矢印の終点にあるデータにも影響がおよびその値が増減するため、政策の対象について考える場合はそのデータに影響を与えるデータのみを考慮することが妥当と考えた。

また、因果グラフにおけるパスの重みが正のときデータ同士も正の相関、負のときに負の相関をとるため、正の場合の始点側のデータを DEA の出力、負の場合の始点側のデータを DEA の入力とすることが妥当と考えた。以上のことから、それらを一つにまとめた csv ファイルを因果探索部分のプログラムのアウトプットとする。因果探索の結果と出力ファイル内のデータの例を図 4.3 に示す。

§ 4.2 楽譜からタブ譜作成のプログラム

本節では、本研究での提案システムにおける DEA 部分の具体的なしくみ、システムの仕様、システムにおいてインプットおよびアウトプットされるデータの形式などを解説する。提案手法における DEA の役割は因果探索によって導かれた政策の対象に対する要因につい

表 4.4: 各都道府県の DMU 内訳

都道府県	市	区	町	村	都道府県	市	区	町	村	都道府県	市	区	町	村	都道府県	市	区	町	村
北海道	35	0	130	20	石川県	11	0	8	0	岡山県	14	4	10	2	愛媛県	11	0	9	0
青森県	10	0	22	8	福井県	9	0	8	0	広島県	13	8	0	9	高知県	11	0	17	6
岩手県	14	0	15	4	山梨県	13	0	8	6	山口県	13	0	0	6	福岡県	27	14	29	2
宮城県	13	5	18	1	長野県	19	0	23	35	徳島県	8	0	15	1	佐賀県	10	0	10	0
秋田県	13	0	9	3	岐阜県	21	0	19	2	香川県	8	0	9	0	長崎県	13	0	8	0
山形県	13	0	19	3	静岡県	21	10	12	0	愛媛県	11	0	9	0	熊本県	13	5	23	8
福島県	13	0	31	15	愛知県	37	16	14	2	高知県	11	0	17	6	大分県	14	0	3	1
茨城県	32	0	10	2	三重県	14	0	15	0	福岡県	27	14	29	2	宮崎県	9	0	14	3
栃木県	14	0	11	0	滋賀県	13	0	6	0	佐賀県	10	0	10	0	鹿児島県	19	0	20	4
群馬県	12	0	15	8	京都府	14	11	10	1	長崎県	13	0	8	0	沖縄県	11	0	11	19
埼玉県	39	10	22	1	大阪府	31	31	9	1	熊本県	13	5	23	8	合計	733	188	727	203
千葉県	36	6	16	1	兵庫県	28	9	12	0	大分県	14	0	3	1					
東京都	26	23	5	8	奈良県	12	0	15	1	宮崎県	9	0	14	3					
神奈川県	16	28	13	1	和歌山県	9	0	20	1	鹿児島県	19	0	20	4					
新潟県	19	8	6	4	鳥取県	4	0	14	1	沖縄県	11	0	11	19					
富山県	10	0	4	1	島根県	8	0	10	1										

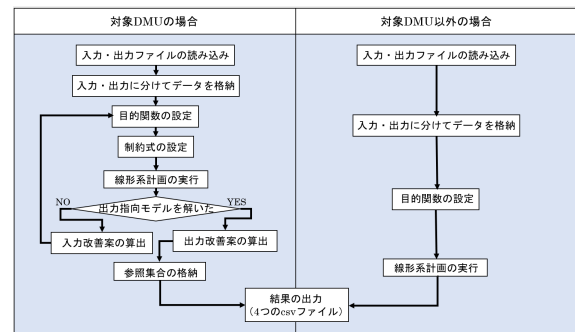


図 4.4: DEA 部分のフロー

て、それらの数値をもとに対象としている自治体の現状を理論的に評価し、その評価を一層高めるために取り組むべき課題を明確にすることによって政策における意思決定の支援をするというものである。

DEA 部分における全体のながれ

DEA 部分では、はじめに 4.1 節での因果探索のアウトプットデータ（図 4.3 参照）を用いて全市区町村を対象に 3.2 節に示した入力指向モデルを解くことによって、各市区町村の評価値を算出する。次に、フロントページにて指定された対象の市区町村に対して、式 (??) および式 (??) を解くことによって、その結果から対象の市区町村における入力・出力の改善案を算出する。

改善案の算出を対象の市区町村のみに限定したのは、全国の市区町村という膨大な DMU を扱う問題において計算量を軽減するためであり、単純な評価値と比較して改善案は対象市区町村以外のものを参考にすることが少ないと考えたからである。

各市区町村に対する評価値の算出

本研究では、47 都道府県に存在する市区町村のうち 773 個の市、188 個の区、727 個の町、203 個の村の合計 1891 個を DMU の集合として、CCR モデルによる DEA 分析を行うことによってそれぞれの DMU における評価値を算出する。各都道府県における DMU の内訳を表 4.4 に示す。因果探索によって対象と因果関係が示された出力データの数を m とすると評価値は以下の線形計画問題を解くことによってもとめられる。

< 評価値の算出式 >

$$\text{minimize} \quad \theta \quad (4.3)$$

$$\text{subject to} \quad \sum_{d=1} y_{id} \lambda_d \geq y_{io} \quad i = 1, 2, \dots, m \quad (4.4)$$

$$- \sum_{d=1} x_{id} \lambda_d + x_{io} \theta \geq 0 \quad i = 1, 2, \dots, m \quad (4.5)$$

$$\lambda \geq 0 \quad (4.6)$$

対象の市区町村に対する入力・出力改善案の算出

フロントページにて対象の市区町村に指定された DMU に対してはすべての入力および出力に対して参照集合のデータをもとにした値の改善案を算出する．参照集合は前述の評価値の算出式において λ_d の値が正の数をとった DMU のみで形成され，参照集合内の DMU の数を A 個，入力の項目数を i 個，出力の項目数を j 個とすると入力および出力の改善案は式 (??) および式 (??) に対して $K = A$ と置くことによってもとめられる．また，4.1 節で述べた通り本研究ではデータに対して前処理を行っているため入力・出力の改善案に関しては逆変換を行い元のデータと同じ単位に戻してから結果を保存している．逆変換の式は以下ようになる．

< 正規化の逆変換 >

$$\iota = \iota' \times 2 \max |\iota| - \max |\iota| \quad (4.7)$$

< robust Z-score の逆変換 >

$$x = \iota \times NIQR + median(x) \quad (4.8)$$

提案手法を実現するプログラム

本研究の提案手法では，DEA の計算に PuLP による線形計画問題のプログラムを使用している．PuLP とは，Python で数理最適化のモデルを記述するためのモジュールであり，PuLP を用いてモデルを記述することによって，混合整数最適化問題を解くことができる．混合整数最適化問題とは，以下のような特徴を持つ数理最適化問題の一種である．

- 連続（実数）変数と整数変数を使って表現される．
- 目的関数と制約条件が 1 次式である．

よって，線形計画問題を用いて定式化することが可能な DEA においても PuLP によるモデルの記述が可能で，PuLP を用いて記述したモデルは同梱される COIN プロジェクトのソルバーである CBC を用いて自動的に解くことができる．問題の定式化を行う回数や算出し，格納するデータの種類の扱う DMU がフロントページで指定された対象の市区町村であるか否かで異なるが，PuLP を用いて最適化問題を解く部分の仕組みは共通である．提案手法の DEA 部分におけるシステムの一連のながれを図 4.4 に示す．

入力・出力のデータは図 4.3 の csv ファイルのように 1 つにまとめられた状態で読み込み，ファイル内の 2 行目に格納されている入力・出力の項目数をもとに 3 行目以降のデータを入力・出力に切り分けることによってそれぞれのデータフレームに格納する．

次に前述の PuLP を用いて入力指向モデルの目的関数，制約式を設定し実行することによって線形計画問題を解き，DMU の評価値を算出する．また，対象 DMU の場合には実行結果をもとに入力改善案を算出する．

その後，同様に PuLP を用いることで出力指向モデルの目的関数，制約式を設定し実行することによって線形計画問題を解く．こちらも対象 DMU の際には実行結果をもとに出力改善案を算出する．

評価値	入力・出力改善案	参照集合・重み（入力）	参照集合・重み（出力）																																																												
<table><tr><th>city_code</th><th>effic_value</th></tr><tr><td>1100</td><td>0.55679</td></tr><tr><td>1202</td><td>0.70558</td></tr><tr><td>1203</td><td>0.83029</td></tr><tr><td>⋮</td><td></td></tr><tr><td>⋮</td><td></td></tr><tr><td>⋮</td><td></td></tr><tr><td>47381</td><td>0.30172</td></tr><tr><td>47382</td><td>0.35638</td></tr></table>	city_code	effic_value	1100	0.55679	1202	0.70558	1203	0.83029	⋮		⋮		⋮		47381	0.30172	47382	0.35638	<table><tr><th>imp_item</th><th>imp_value</th></tr><tr><td>AAA</td><td>0</td></tr><tr><td>BBB</td><td>13.244</td></tr><tr><td>CCC</td><td>0.86798</td></tr><tr><td>DDD</td><td>5.68598</td></tr><tr><td>EEE</td><td>3.69084</td></tr><tr><td>FFF</td><td>26.6893</td></tr><tr><td>GGG</td><td>6.20473</td></tr><tr><td>HHH</td><td>42026.3</td></tr></table>	imp_item	imp_value	AAA	0	BBB	13.244	CCC	0.86798	DDD	5.68598	EEE	3.69084	FFF	26.6893	GGG	6.20473	HHH	42026.3	<table><tr><th>city_code</th><th>weight_in</th></tr><tr><td>24303</td><td>0.0611</td></tr><tr><td>33663</td><td>0.17123</td></tr><tr><td>34203</td><td>0.01215</td></tr><tr><td>39424</td><td>0.02395</td></tr><tr><td>44462</td><td>0.32172</td></tr></table>	city_code	weight_in	24303	0.0611	33663	0.17123	34203	0.01215	39424	0.02395	44462	0.32172	<table><tr><th>city_code</th><th>weight_out</th></tr><tr><td>24303</td><td>0.13164</td></tr><tr><td>33663</td><td>0.3689</td></tr><tr><td>34203</td><td>0.02617</td></tr><tr><td>39424</td><td>0.05159</td></tr><tr><td>44462</td><td>0.69311</td></tr></table>	city_code	weight_out	24303	0.13164	33663	0.3689	34203	0.02617	39424	0.05159	44462	0.69311
city_code	effic_value																																																														
1100	0.55679																																																														
1202	0.70558																																																														
1203	0.83029																																																														
⋮																																																															
⋮																																																															
⋮																																																															
47381	0.30172																																																														
47382	0.35638																																																														
imp_item	imp_value																																																														
AAA	0																																																														
BBB	13.244																																																														
CCC	0.86798																																																														
DDD	5.68598																																																														
EEE	3.69084																																																														
FFF	26.6893																																																														
GGG	6.20473																																																														
HHH	42026.3																																																														
city_code	weight_in																																																														
24303	0.0611																																																														
33663	0.17123																																																														
34203	0.01215																																																														
39424	0.02395																																																														
44462	0.32172																																																														
city_code	weight_out																																																														
24303	0.13164																																																														
33663	0.3689																																																														
34203	0.02617																																																														
39424	0.05159																																																														
44462	0.69311																																																														

図 4.5: DEA のアウトプットの例

加えて、入力・出力の改善案を算出する際に用いた参照集合内の市区町村名およびそれらの市区町村にかかるウェイトに関してもリストを作成して格納する。最後にこれらの結果をもとに以下に示す 4 つの csv ファイルをアウトプットする。アウトプットする 4 つの csv ファイルの例を図 4.5 に示す。

- 全市区町村を対象にしたそれぞれの市区町村に対する評価値
- 対象の市区町村における入力・出力の各項目に対する改善案
- 対象の市区町村における入力の改善案を算出した際の参照集合とそれぞれのウェイト
- 対象の市区町村における出力の改善案を算出した際の参照集合とそれぞれのウェイト

DEA による分析結果を以上のような 4 つの csv ファイルに分割したのはそれぞれのデータにおける対象の適用範囲や役割が異なり、4.3 節で後述する結果の提示の際に別の提示方法によってユーザにフィードバックされるためである。具体的には、1 つ目の csv ファイルが全ての市区町村に対してそれぞれにもとめられる評価値であり、各行に市区町村ごとの値が格納されているのに対して、その他の csv ファイルのデータはフロントページで指定された 1 つの市区町村に対する各項目のデータが格納されているという違いがある。ゆえに、システムにおけるデータの読み込みおよび取り扱いの簡便化のために個別のファイルでアウトプットすることとした。

§ 4.3 楽曲からタブ譜作成システムの開発

本節では、4.1 節、4.2 節で述べたデータの選別・分析手法による結果に基づいて 3.3 節で解説した Folium による Web-GIS 作成を行うことで分析結果をユーザに提示する提案手法における結果のフィードバック部分について解説する。

また、これまでに 4 章の中で提示した技術のそれぞれがどの部分に組み込まれているかについて再度整理しながら、提案手法を組み込んだシステムの全体の流れを解説する。加

えて、本研究の提案手法における WebGIS を用いたデータフュージョンの方法をシステム内の実際の仕様に言及しながら示す。

Web-GIS アプリケーションの作成

図 4.5 に示した 4 つ csv 内のデータをもとに Web-GIS による結果のフィードバックを行うアプリケーションである EBPM-GIS を作成した。各 csv ファイルが EBPM-GIS のどの部分の実装に使用されているかを図 4.6 に示す。

EBPM-GIS におけるベースマップは cartodbpositron および地理院タイルの 2 種類であり、cartodbpositron は白色でマーカーが視認しやすいという理由、地理院タイルは等高線や色分けによって地形が表されており、土地の地理的状況が視覚的に理解できるという理由で採用した。

各市区町村における評価値は矢印のアイコンを用いたマーカーで表示し、矢印の向きと色は三種類に分けて実装した。基準は評価値の大きさで、0.33 未満を青の下向き、0.33 以上 0.66 未満を黄の横向き、0.66 以上を赤の上向きとした。また、対象の市区町村および参照集合の市区町村に関してはアイコンの色も変更し、対象を赤、参照集合を青とした。

EBPM-GIS では、データ分析に用いた市区町村すべてに対してマーカーを立てるため、非常に多くのマーカーが画面上に表示される。そのため、視認性および処理速度の向上を目的として前述の矢印の種類ごとに別のレイヤに分けて実装した。各レイヤの切り替えは画面右上のレイヤコントロールで行うことができる。初期画面では全ての矢印が表示された状態となっている。

EBPM-GIS 上でマーカーをクリックするとポップアップが出現し、DEA における結果がテキストで表示される。対象の市区町村に関しては市区町村名、評価値、入力・出力の改善案、参照集合とその重みがポップアップ上に表示される。その他の市区町村においては市区町村名および評価値の表示のみである。

また、因果探索によって対象のデータと地理情報データの間に関係が示された場合、全国に分布するその施設の位置全てをマーカークラスタによって表示する。このマーカークラスタも評価値のマーカーと同様にベースマップから切り離された別のレイヤ上にプロットされており、レイヤコントロールを用いて表示・非表示の切り替えが可能である。

提案手法全体の流れ

1 章で示した政策の対象となる問題に関する要因の複雑性という課題に対して、4 章で示した各手法を統合した課題解決のための提案手法全体の流れの説明を行う。また、提案システム全体のフローを図 4.7 に示す。

Step 1: 政策の対象決定・データ選択

フロントページにてサーバ上に存在する統計データ全ての項目名を公開し、全国の市区町村に対する項目ごとのデータを参照できるようにする。政策に関係した問題を抱えるユーザはその問題に関係のあるデータの項目名および自身が取り扱いたい市区町村名を各 1 つずつ選択形式でシステムに入力する。選択されたデータは次のフローへ送信される。4.1 節の前半にて解説した部分がこれにあたる。

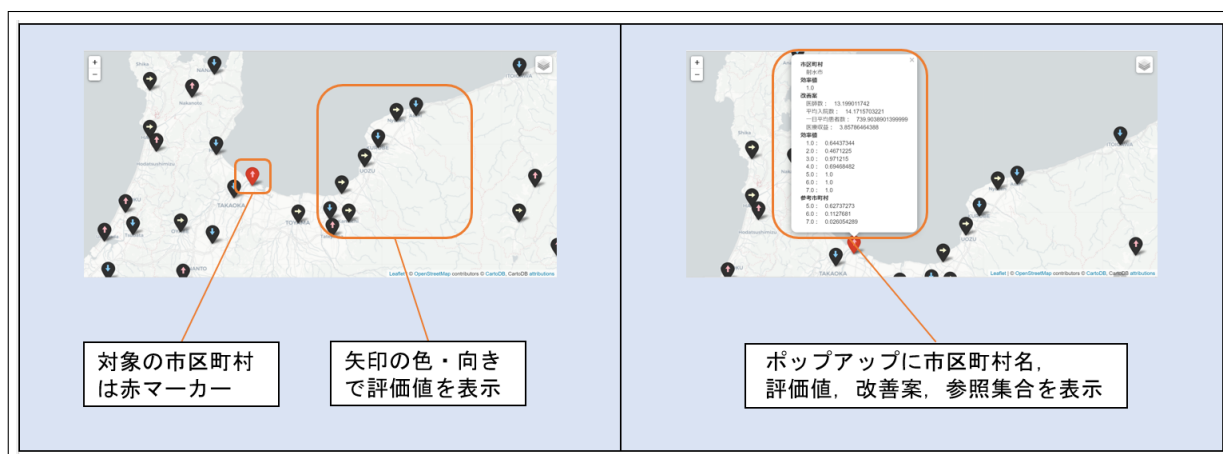


図 4.6: EBPM-GIS の実装

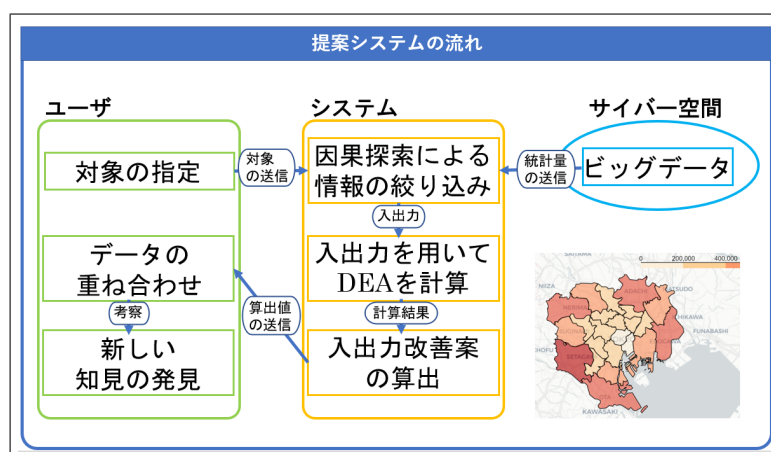


図 4.7: 提案システム

Step 2: 統計データの選別・入出力への振り分け

Step1 で選択されたデータを受信し、受信したデータに含まれる対象のデータ項目名をターゲットとしてサーバ上に存在する全ての統計データを用いた Direct-LiNGAM による因果探索を行い、対象のデータ項目と関係のあるデータのみを選別する。

また、選別されたデータのうち、対象のデータ項目に対して因果関係が向いているデータを入力、対象のデータ項目から因果関係が向いているデータを出力とすることで選別されたデータを DEA における入力・出力データに振り分ける。選別されたデータは入力・出力の項目数とともに 1 つの csv ファイルにまとめて排出する。4.1 節の後半にて解説した部分がこれにあたる。

Step 3: DEA による評価値と入力・出力改善案の算出

Step2 にて作成された csv ファイル内の入力・出力の項目数をもとに入力・出力データを参照し、それらを対象とした DEA (CCR モデル) を解くことによって全市区町村を対象にそれぞれの評価値を算出する。

また、Step1 のフロントページにて選択された市区町村に対して DEA の入力指向モデルおよび出力指向モデルを用いた入力・出力改善案の式を適用することによって対象の市区町村における入力および出力それぞれに対する改善案を算出する。

以上の操作の結果もとめられた全市区町村に対する評価値、対象の市区町村に対する入力・出力の改善案に加えて改善案を算出する際に用いた参照集合とその重み（入力指向モデル、出力指向モデルでもとめられるそれぞれの値）をデータの種類ごとに4つの csv ファイルにして排出する。4.2 節のにて解説した部分がこれにあたる。

Step 4: EBPM-GIS の作成・データの重ね合わせ

Step3 にて作成された4つ csv ファイルをもとに結果の表示および地理情報データとの重ね合わせによるデータフュージョンによって政策決定を支援することを目的とした EBPM-GIS を作成・表示する。

4つの csv ファイルのデータのうち、全市区町村に対する評価値のデータは GIS 上にマーカーとして表示し、その値の大小によって3つのレイヤに分ける。その他のデータはマーカーのポップアップ内にテキストとして表示する。

また、Step2 にて対象のデータ項目と因果関係が示されたデータのうち、施設の場所など地理的な特徴を持つデータに関してはそれらのデータ単体でマーカークラスタとして表示し、Step3 での分析結果と重ね合わせることによってデータフュージョンを行う。4.3 節の前半にて解説した部分がこれにあたる。

システム上でこれら4つの Step の処理が行われることによって出力装置には最終的に EBPM-GIS が出力されることになる。つまり、ユーザサイドから見た場合、フロントページにて対象のデータ項目および市区町村を入力し実行すると EBPM-GIS が表示されるという画面遷移だけが提示される。

また、ユーザは出力装置に提示される EBPM-GIS を通して自身が対象としたいデータ項目に因果関係のあるデータにはどのようなものがあるか、それらを用いて DEA を行った結果の評価値や入出力の改善案、改善のために参考とすべき市区町村とその重みはどのようなであるか、対象としたいデータと因果関係を持つ施設はどのように分布しているかを受け取ることができる。

ユーザは、これらのデータに対して切り替えや重ね合わせを用いることで考察を行い、政策における意思決定を行う。どのように考察するかはユーザの自由であるが、現段階では評価値によって自身の市区町村の現状を知り、周辺の市区町村との差などを「感覚的に理解する。改善案や優れた市区町村をもとに目標を明確化する。そのうえで、重ね合わせ等によって政策決定における新たな知見を得る」という方法を想定している。

数値実験並びに考察

§ 5.1 数値実験の概要

本研究の数値実験としてシステムの有用性の検証、科目推薦の有効性を行う。システムの有用性の検証では実際にシステムを使用してもらい、アンケートに答えてもらう。アンケートの項目は全部で 10 個あり、その 10 個には必ず答えてもらう。また、アンケートと同時にコメントを記入できる欄を設けておき、自由にコメントをできるようにする。このアンケートを持って本システムの有用性の検証を行う。アンケート項目は表 5.1 に示す。以上のコメントを 5 段階のリッカート尺度で評価してもらう。リッカート尺度とは、あるトピックに対して、多段階の選択肢を用いたアンケートを取り、回答者がどの程度同意するか測定する手法である。今回のアンケートでは 5 段階のうち、1 を全く満足していない、2 をあまり満足していない、3 をどちらでもない、4 をやや満足している、5 を非常に満足しているといったようにアンケートに答えてもらう。表 5.1 を見てわかる通り、アンケート項目全体を通して、基本的にはシステムの使用感に関する質問を多くしてある。本来なら本システムを使用し、成績（GPA）が向上するのを確認することでシステムの有効性を検証すべきだが、成績向上を確認するには開発したシステムを最低でも半期使用してもらわないと成績が向上したかを確認することができない。そのため、本研究ではアンケート調査を用いてシステムの有効性を示す。調査の対象は同研究室の修士 1 年、学部 4 年、3 年生の合計 11 人に実際に開発したシステムを使用してもらい、アンケートを答えてもらった。実際に使用してもらうにあたり、システムの使用手順について説明を行い、実際に使用してもらう。手順は以下に記してある通り、新規登録から推薦科目の選択、教材のレビュー、ログアウトまでの一連の流れを説明した。

取り上げるモデルケース

本実験では、問題を抱える地方自治体のモデルケースとして、富山県立大学工学キャンパスの所在である富山県射水市を取り上げる。当市は富山県の沿岸部中央付近に位置し、土地面積 109.43km^2 、人口約 90,000 人を有する市である。

また、現在の市域に至ったのは 2005 年 11 月で当時の新湊市、小杉町、大門町、大島町、下村の 1 市 3 町 1 村が合併することで射水市となった。そのため、富山県に占める土地面積は比較的広大で前述の人口は富山市、高岡市に次ぐ県内 3 位の値となっている。

以上のような特徴を持つ当市であるが、当市が抱える問題として最も深刻なものに人口減少問題が挙げられる [21]。当市の人口は 2005 年にピークを迎えて以来、年々減少しており、特に老年人口の増加に反比例して若年人口が減少することで少子高齢化に歯止めが利

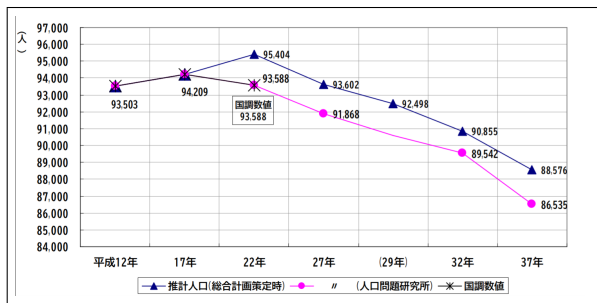


図 5.1: 射水市の推計人口 [22]

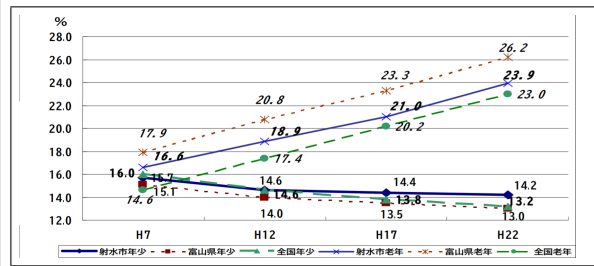


図 5.2: 年少・老年人口割合推移 [22]

かない状況となっている。20 歳から 24 歳までの若い世代の人口が他市町村へ流出，とりわけ女性の転出の増加が原因の 1 つと考えられる。2010 年の国勢調査に基づく当市の推計人口の推移を図 5.1，年少（15 歳未満）・老年（65 歳以上）人口割合の推移を図 5.2 に示す。

国立社会保障・人口問題研究所が作成した当市における 2025 年の推計人口は 86,535 人であり，統計上最多である 2005 年の 94,209 人から約 10% 減少した値となっている。また，1995 年から 2010 年にかけて老年人口割合が約 8% 増加し，年少人口が約 3% 減少していることから前述の少子高齢化が統計からも読み取れる。

以上の問題のうち，少子化に対して 2005 年に「射水市次世代育成支援行動計画」を策定，2007 年には「射水市子ども条例」を制定した。また，これらの理念に基づき，子ども医療費助成の中学校 3 年生までの拡大や第 3 子以降保育料の無料化，様々なニーズに応じた保育サービスの実施，放課後児童クラブの拡充といった福祉，保健サービスを始め，子どもの学び・育ちを支えるための学校教育の充実，子どもに関する相談体制の整備などの取り組みを行った。

これらの取り組みは 2005 年から 2022 年現在も継続して行われており 2021 年末における人口が 91,783 人と発表されていることから推計と比較して人口減少のペースは緩やかになっていると考えられる。ただ，いまだ人口が減少傾向であることは変わらず，2020 年 3 月に策定され，2021 年 3 月に改訂された「第 2 期射水市まち・ひと・しごと創生総合戦略」[] においても戦略全体の方向性として人口減少に対する対策が言及されている。このことから，本研究における提案手法の有効性評価として，「今後，射水市の人口減少を食い止めるためにはこういった政策を行うべきか」を対象とした数値実験を行うこととする。

数値実験のながれ

前述の問題に対してデータベース内のデータ項目のうち，最も関係があると考えられる「年少人口 [割合]」をターゲットとして提案手法による分析を行う。また，対象とする市区町村には射水市を選択する。

はじめに，これらの情報をフロントページを用いてシステムに送信し，「年少人口 [割合]」をターゲット，観測変数をデータベース上に存在する全てのデータとして Direct-LiNGAM による因果探索を行う。因果探索の結果は図 4.3 のような csv 形式で示す。

次に，その csv に格納されているデータを入力・出力とし，対象 DMU を射水市，DMU の集合をデータベース上に存在する全ての市区町村とする DEA を行う。分析結果については評価値，入力・出力の改善案，入力指向モデルおよび出力指向モデルにおける参照集合に属する市区町村名とそれらに対する重みを項目ごとに分けて 4 つの csv で示す。

最後に、それらの csv ファイルをもとに作成された EBPM-GIS における射水市のポップアップおよび参照集合に属する市区町村、射水市周辺の市区町村のポップアップを示し、それらの結果を比較する。また、射水市と前述の市区町村の政策、特に因果探索によって因果関係が示されたデータに関するものを比べることで射水市が年少人口を増大させ、少子高齢化を食い止めていくためには今後どのような政策に取り組んでいくべきなのかを考察することとする。

§ 5.2 実験結果と考察

5.1 節で述べた通り、富山県射水市が抱える少子高齢化問題に対して今後どのような政策に取り組んでいけばよいかを模索するというモデルケースを想定して本研究の提案手法を用いた意思決定の支援を行うという数値実験を行った。その結果を 4.3 節で述べた 4 つの Step のうち、Step2 から Step4 における実行結果に基づいて示す。

統計データの選別・入出力への振り分け

因果探索の部分では、少子高齢化問題を解決するにあたってその値の増加が必要であると考えられるデータ項目である「年少人口 [割合]」をターゲットとしてデータベース内のデータ間における因果関係を Direct-LiNGAM を用いてもとめる。「年少人口 [割合]」とは、各市区町村における人口全体のうち、年少（15 歳未満）の人口が占める割合のことである。

Direct-LiNGAM によってもとめられたパス係数行列の中で「年少人口 [割合]」に対するパス係数が 0 ではないものの項目名とそのパス係数を表 5.1 に示す。また、結果に基づく DEA の入力・出力の項目を表 5.2 に示す。

入力・出力の項目数はパス係数の大きいものから老年人口、住宅用地平均取引価格、警察費・消防費を含む 6 つの入力と生産年齢人口、施設数 [空港]、教育費の 3 つの出力となった。本実験で用いたデータベースには老年人口、生産年齢人口、年少人口と 3 つの区分における人口割合が含まれる。これらのデータは割合のためお互いに作用しあうということから老年人口、生産年齢人口から因果関係が示されたのは妥当と考えることができる。

そのうえで、老年人口が入力、生産年齢人口が出力に振り分けられたのは年少人口の親世代が生産年齢人口にあたるためと考えられる。また、生産年齢人口を増やそうと考えた際に住宅用の土地の取引価格が低いほうがよいというのは容易に想像できる。教育費が出力に振り分けられたのも同様の理由が考えられ、安心して育児・教育に取り組める環境がある場合に年少人口が増加すると考えられる。

逆に空港の施設数や警察費・消防費から因果関係が示されたこと、企業数が入力に振り分けられたことなどは直接的な要因が想像しづらく、こういった項目が分析によって導き出されることで新たな発見につながるこれが本研究における有意性の 1 つといえる。

DEA による評価値と入力・出力改善案の算出

DEA の部分では、前述の因果探索における結果を入力・出力として DEA を行い、射水市における評価値、入力・出力改善案、参照集合に属する市区町村名とそれらにかかるウェイトおよび全国の市区町村における評価値を算出する。DEA に用いた入力・出力が格納さ

データ項目	パス係数	データ項目	パス係数
施設数 [空港]	0.059	企業数	-0.006
衛生費	-0.019	商工費	-0.024
警察・消防費	-0.038	教育費	0.017
住宅用地平均取引価格	-0.043	生産年齢人口	0.249
老年人口	-0.559		

入力	出力
住宅用土地平均取引価格	生産年齢人口
警察費・消防費	教育費
商工費	施設数 [空港]
衛生費	
企業数	
老年人口	

[illegible]

れている csv ファイルの中身を 5.3 に示す。また、射水市の合計 10 項目ある入力・出力のそれぞれに対する元のデータと改善案との比較を 5.3、入力指向モデルおよび出力指向モデルにおける参照集合に属する市区町村名とそれらにかかるウェイトを 5.4 に示す。

本実験で対象としている年少人口は増加が望まれており、正の因果を持つ項目の数値の増加と負の因果を持つ項目の数値の減少によって年少人口の数値の増加が望めるので妥当な改善値であるといえる。逆に、犯罪率などを対象とする際は、対象データ自体を現状させることが望まれるので、DEAの分子と分母を逆にする処理が必要である。

また、参照集合には山形県東置賜郡川西町、長野県北佐久郡御代田町、滋賀県米原市、同県蒲生郡日野町が含まれており、川西町がとりわけ高くなっている。これらのことから、川西町をはじめとした参照集合に属する市町村での取り組みなどを参考にしつつ、目的別の歳出額の調整や住宅用の土地取引に関する補助金などのしくみを作っていくことが年少人口の増加につながると考察できる。

表 5.3: 元データと改善案の比較

	元データ	改善案
企業数	3075	-27769.854 (社)
衛生費	10.08	-2.954 (%)
商工費	2.92	-1.647 (%)
警察費・消防費	2.88	-1.470 (%)
住宅土地平均取引価格	19507	-368560.385 (円)
老年人口	28.6	22.821 (%)
施設数 [空港]	0	0.143 (箇所)
教育費	12.5	32.502 (%)
生産年齢人口	57.5	66.692 (%)

表 5.4: 参照の市区町村とウェイト

入力指向モデル		出力指向モデル	
参照市区町村	ウェイト	参照市区町村	ウェイト
山形県東置賜郡川西町	0.268	山形県東置賜郡川西町	0.315
長野県北佐久郡御代田町	0.197	長野県北佐久郡御代田町	0.231
滋賀県米原市	0.108	滋賀県米原市	0.127
滋賀県蒲生郡日野町	0.158	滋賀県蒲生郡日野町	0.186

おわりに

本研究では、市区町村単位の自治体における EBPM の支援を目的として、政策に関する意思決定のためのデータ収集および分析を行う手法を提案した。まず、不特定多数のオープンデータを用いることで、政策の対象によって収集するデータが偏ることを防ぎ、その中から LiNGAM によって対象と因果関係があるデータを選別した。

次に、因果関係が示されたデータのうち、対象のデータに対してパスを持つデータをパスの正負に着目しながら二つのグループに分けることによって、DEA の入力・出力とした。そして、DEA の CCR モデルを用いて各自治体における評価値および入力・出力の改善案を算出することによって、対象の自治体における現状を評価し、現状以上に優れた運営を行うためには入力・出力をどのように変化させればよいかを分析した。また、参照集合に属する市区町村とそれらにかかるウェイトを提示することで入力・出力を改善するうえで参考とすべき自治体を示した。

最後に、それらの結果に基づいてデータのフィードバックを行う EBPM-GIS を作成し、各自治体における評価値と対象の自治体に関しては、入力・出力の改善値、参照集合に属する市区町村名とそれぞれに対する重みを地図上にプロットした。また、LiNGAM によって因果関係が示されたデータのうち、位置情報に紐づくものに関してはマーカークラスターを用いて位置情報をプロットしたレイヤーを作成することで DEA による分析結果との重ね合わせを可能にし、データフュージョンによるさらなる知見の発見を可能にした。

今後の課題としては、データベース内のデータの拡充、因果探索および DEA におけるモデルや分析手法の深化などが挙げられる。本研究で取り上げた課題である政策決定における問題の複雑性の性質上、分析に用いるデータの集合は多種多様かつ多量であるほど課題解決に有意義な結果が得られることが考えられる。そのため、実際の政策決定で用いる際には本研究で扱ったデータベースからデータ量を大幅に増加させる必要があると考える。

また、本研究で用いた分析手法は因果探索や DEA における基本的なモデルであるため、多様な問題やデータに対応するためにはそれらによりフィットした分析手法やモデルについての議論が必要であると考えられる。特に、因果探索においてはどの程度のパス係数までを有意として因果関係を認めるかなど今後発展の余地が残されている。

今後の展望としては、本研究で扱ったような自治体における政策決定への適用だけでなく、社会活動を行う上で頻繁に遭遇する対象が複雑な問題やデータ分析に際してデータ群の全容が確定していない問題に対して本研究における提案手法を応用していくことも考えられる。

謝辞

本研究を遂行するにあたり，多大なご指導と終始懇切丁寧なご鞭撻を賜った富山県立大学工学部情報システム工学科情報基盤工学講座の奥原浩之教授，António Oliveira Nzinga René 講師に深甚な謝意を表します．また，システム開発および数値実験にあたり，ご助力いただいた富山県立大学電子・情報工学科３年生の島部達哉氏に感謝の意を表します．最後になりましたが，多大な協力をしていただいた研究室の同輩諸氏に感謝致します．

2022 年 2 月

山内拓海

参考文献

- [1] 水野 創太, 一ノ瀬修吾, 白松俊, 北原鉄朗, ”演奏未経験者のためのスマートフォンセンサーを用いた即興合奏支援システムの試作”, The 31st Annual Conference of the Japanese Society for Artificial Intelligence, 2017.
- [2] 長谷川隆, 西本卓也, 小野順貴, 嵯峨山茂樹, ”楽譜情報からの作曲家らしさ認識のための音楽特徴量の提案”, 情報処理学会論文誌 Vol.53 No.3 1204-1215 (Mar. 2012).
- [3] 平賀悠介, 大石康智, 武田一哉, ”主観評価に基づく楽曲間類似度算出モデル”, 情報処理学会研究報告, Vol.2009-MUS-81 No.2, 2009/7/29.
- [4] 張夢堯, 成凱, ”鼻歌検索のための音楽特徴量の抽出と評価”, DEIM Forum 2022 C21-4.
- [5] 矢澤一樹, 糸山克寿, 奥乃博, ”ギター演奏音からの難易度調整可能なタブ譜自動生成システム”, 情報処理学会第 76 回全国大会, 5R-5.
- [6] 佐藤主光, ”税財政分野における EBPM の基礎と活用”, 閲覧日 2022-02-08, https://www.ipp.hit-u.ac.jp/satom/lecture/localfinance/2019_local_note07.
- [7] 内閣府, ”内閣府における EBPM への取組”, 閲覧日 2022-02-08, <https://www.cao.go.jp/others/kichou/ebpm/ebpm.html>.
- [8] esri ジャパン, ”GIS (地理情報システム) とは”, 閲覧日 2022-02-08, <https://www.esri.jp/getting-started/what-is-gis/>.
- [9] 国土交通省国土地理院, ”基盤地図情報の利活用事例集”, 閲覧日 2022-02-08, <https://www.gsi.go.jp/common/000062939>.
- [10] esri ジャパン, ”東日本大震災対応における政策形成支援に GIS を活用”, 閲覧日 2022-02-08, <https://www.esri.jp/industries/case-studies/35859/>.
- [11] 田中貴宏, 佐土原聡, ”都市化ポテンシャルマップと二次草原潜在生育地マップの重ね合わせによる二次草原消失の危険性の評価: 一福島県旧原町市域を対象として”, 環境情報科学論文集, Vol. 23, pp. 191-196, 2009.
- [12] 坪井利樹, 西田佳史, 持丸正明, 河内まき子, 山中龍宏, 溝口博, ”身体地図情報システム”, 日本知能情報ファジィ学会誌, Vol. 20, No. 2, pp. 155-163, 2008.
- [13] 杉原豪, 塚井誠人, ”統計的因果探索による社会基盤整備のストック効果の検証”, 土木学会論文集 D3 (土木計画学), Vol. 75, no.6, pp. 583-589, 2020.
- [14] Dentsu Digital Tech Blog, ”Google Colab で統計的因果探索手法 LiNGAM を動かしてみた”, 閲覧日 2022-02-08, https://note.com/dd_techblog/n/nc8302f55c775.

- [15] 藤井秀幸, 傅靖, 小林里佳子, ”データ包絡分析を用いたふるさと納税の戦略提案-K市のふるさと納税への適用事例-”, 日本経営工学会論文誌, Vol. 71, No. 4, pp. 149-172, 2021.
- [16] 刀根薫, ”包絡分析法 DEA”, 日本ファジィ学会誌, Vol. 8, No. 1, pp. 11-14, 1996.
- [17] 金成賢作, 篠原正明, ”DEA における入力指向と出力指向の比較 (その 1) ”, 日本大学生産工学部第 42 回学術講演会, 2009.
- [18] 日本オペレーション・リサーチ, ”第 4 章 包絡分析-入力と出力と”, 閲覧日 2022-02-08, <http://www2.econ.tohoku.ac.jp/ksuzuki/teaching/2006/ch4>.
- [19] pork_steak, ”folium 事始め”, 閲覧日 2022-02-08, https://qiita.com/pork_steak/items/f551fa09794831100faa.
- [20] 保母敏行ほか, ”日本分析学会における標準物質の開発”, 日本分析化学会誌, vol. 57, No. 6, pp. 363-392, 2008.
- [21] 射水市役所, ”総合戦略-射水市”, 閲覧日 2022-02-08, <https://www.city.imizu.toyama.jp/appupload/EDIT/054/054185>.
- [22] 射水市役所, ”共通課題-射水市”, 閲覧日 2022-02-08, <https://www.city.imizu.toyama.jp/appupload/EDIT/024/024383>.

