

令和 2 年 6 月 16 日現在

機関番号：14301

研究種目：基盤研究(B)（一般）

研究期間：2016～2019

課題番号：16H02789

研究課題名（和文）データベクトル間の関連を考慮した多ドメインデータの多変量解析

研究課題名（英文）Multivariate analysis of multi-domain data considering the association between data vectors

研究代表者

下平 英寿（Shimodaira, Hidetoshi）

京都大学・情報学研究科・教授

研究者番号：00290867

交付決定額（研究期間全体）：（直接経費） 13,000,000円

研究成果の概要（和文）：画像、タグ、文書等の様々な種類（ドメインと呼ぶ）の情報源から得られるデータベクトルと、データベクトル間の関連性の強さに関する情報が得られているデータを、多ドメイン関連性データと呼ぶ。従来の多変量解析ではベクトルが1対1対応するものを扱っていたため、柔軟なデータ構造を表すことができなかった。本研究ではベクトルの関連性をグラフ（ネットワーク）で表現して、そのグラフ構造をなるべく保存するように次元削減を行う情報統合の方法を提案・発展させた。

研究成果の学術的意義や社会的意義

関連性データのグラフ構造をなるべく保存するようにデータベクトルを変換することをグラフ埋め込みという。正準相関分析など従来の多変量解析を一般化したグラフ埋め込み手法を提案し、画像と単語の相互検索などのタスクで有効性を確認した。ニューラルネットワークによる非線形変換を用いたグラフ埋め込み法を提案し、さらに外れ値の影響を軽減するロバスト化を行った。ベクトル間の内積とそれを発展させたニューラルネットワークモデルによって表現できる類似度関数のクラスを明らかにした。

研究成果の概要（英文）：Multi-domain association data consists of data vectors from various types of sources (called domains), such as images, tags, documents, etc., and the strength of associations between data vectors. Conventional multivariate analysis has dealt with one-to-one vector correspondence, and therefore, it cannot represent flexible data structures. In this study, we describe the relationship between vectors as a graph (network). Then, we have proposed and developed methods of information integration via dimensionality reduction, which preserves the graph structure as much as possible.

研究分野：統計学・機械学習

キーワード：多変量解析 パターン認識 グラフ埋め込み 次元削減 分散表現 ニューラルネットワーク マルチモーダル 自然言語処理

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

## 様式 C-19、F-19-1、Z-19 (共通)

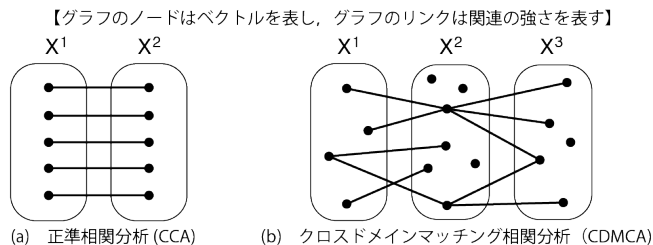
### 1. 研究開始当初の背景

画像、タグ、文書、音声、位置・時刻などは異なる種類の情報であり、これらを「ドメイン」と呼ぶ（ビュー、モーダルなどとも呼ばれる）。たとえばソーシャルメディアでは画像、タグ、地理情報等の様々な種類（ドメイン）のデータが得られる。このように、複数のドメインがあつて、それぞれのドメインの要素がデータベクトルで表されるようなデータは **multi-domain, cross-domain, multi-modal, multi-view** データなどと呼ばれ、その機械学習が研究されている。そして多変量解析、とくに正準相関分析を発展させて多ドメインデータの次元削減によって情報統合を行うものがパターン認識、ビジョン、自然言語処理の分野でさかんに利用、研究されている。

正準相関分析 (CCA; Hotelling 1936) では二つのドメインを扱う。

CCA を多ドメインへ拡張した多変量解析は **Multiset Canonical Correlation Analysis (MCCA; Kettenring 1971)** と呼ばれる。正準相関分析 (CCA や MCCA) など従来の多変量解析では各ドメインのサンプルサイズが同一でベクトルの対応が 1 対 1 (**one-to-one**) 対応

のものを扱っていた。しかし、これではソーシャルメディアのような多様な対応関係を柔軟に表現することはできない。このような背景で、本研究の準備段階で柔軟な多変量解析の方法論としてクロスドメインマッチング相関分析 (**Cross-Domain Matching Correlation Analysis; CDMCA**) を提案した (Shimodaira 2016; *Neural Networks*, Vol. 75, pp. 126-140)。CDMCA では、各ドメインのサンプルサイズは一般に異なり、ベクトルの対応も多対多 (**many-to-many**) 対応のものが扱える。CDMCA は CCA や主成分分析 (PCA) などの多変量解析を特殊な場合として含む一般化である。



データベクトル間の対応関係の構造はグラフ(またはネットワーク)として表現することができ、データベクトルが各ノードに割り当てられていると考える。ベクトル間の関連性の強さをリンクの重みとして任意に設定することで、柔軟なデータ解析を行える。このようなデータを多ドメイン関連性データと呼ぶことにする。CDMCA では各ドメインのデータベクトルを線形変換して共通の低次元空間におけるベクトル表現を得る。もとのグラフ構造をなるべく保存するような低次元表現をグラフ埋め込みとよぶが、CDMCA はその一種である。

### 2. 研究の目的

多ドメイン関連性データを次元削減するための新しい多変量解析である CDMCA を発展させ、多変量解析の方法論を展開することが本研究の目的である。画像認識や自然言語処理などの応用にも取り組み、そこでの知見もふまえて理論と手法の研究にとりくむ。多ドメイン関連性データの次元削減をグラフ埋め込みとみなして、ロバスト化や非線形化など、理論と手法の拡張をおこなう。

### 3. 研究の方法

つぎのようなアプローチで研究をすすめた。

- (1) CDMCA の計算はある種の固有値、固有ベクトル問題になる。データサイズが数十万、データベクトルの次元が数千程度では、スパース行列ライブラリを使う程度の工夫で実装できる。さらに大規模な問題では、確率的アルゴリズム (Halko et al., SIAM Rev. 2011) の利用を検討する。
- (2) データがなんらかのノイズ等の影響で外れ値となることがある。例えば、ソーシャルメディアで画像-タグの間違ったリンクが与えられた場合、それを外れ値としてダウンウェイトしたデータ解析を行いたい。このように外れ値にたいしてロバストな手法を検討する。
- (3) 自然言語処理における単語埋め込みは単語と文脈の関連性を保存するような単語のベクトル表現を得る手法であるが、これはグラフ埋め込みの一種と考えられる。ソーシャルメディアは単語と画像の関連性が与えられるので、グラフ埋め込みによって単語や画像のベクトル表現を得ることができる。このように関連する分野の応用に積極的にとりくむ。
- (4) リサンプリングによって推定結果の信頼度を計算するための方法論を検討する。

- (5) CDMCA はデータベクトルを線形変換して次元削減をおこない、共通空間のベクトル表現を得るが、これを非線形変換に置き換えた一般化を検討する。

#### 4. 研究成果

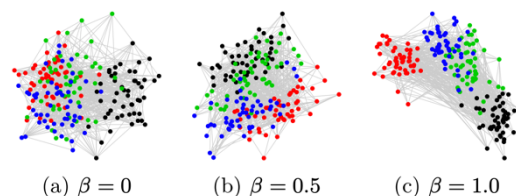
主な成果について概要を述べる。

(1) 多言語単語埋め込み (Oshikiri et al., 2016; ACL) や画像と単語の同時埋め込み (Fukui et al., 2016; ICIP) に実際に CDMCA を適用してその有効性を確認した。Fukui et al. (2017; TextGraphs) では、写真共有サイト Flickr のタグ付けされた画像データから「画像-タグ」のリンク、Wikipedia の文書データから「単語-文脈」のリンクを入手し、単語ベクトルと画像ベクトルを 500 次元の共通空間に埋め込んで、単語と画像の加減算を含む相互検索を実現した。大規模データでは CDMCA の実装に確率的アルゴリズムを利用して有効性を確認した。



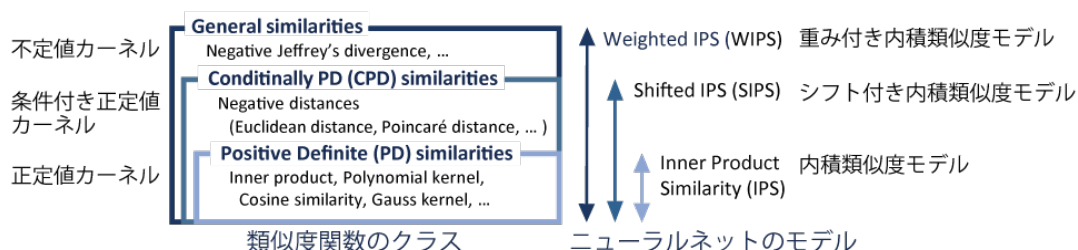
(2) CDMCA の線形変換をニューラルネットワークによる非線形変換に置き換えて、グラフ埋め込みによる次元削減を一般化する方法を Okuno et al. (2018; ICML) で提案し、Probabilistic Multi-view Graph Embedding (PMvGE) と呼んだ。ただし確率的な生成モデルを導入して、リンク重みが例えばポアソン分布やベルヌーイ分布に従うと仮定してモデル化した。PMvGE は、CDMCA だけでなく言語モデルによる単語ベクトルなど様々なモデルを (近似的に) 特殊な場合として含むモデルになっている。このような生成モデルの最尤推定は CDMCA で必要とされた制約条件が不要となり、ニューラルネットワークの学習で用いられる確率的勾配降下法 (SGD) や単語埋め込みで用いられる negative sampling などの最適化技法によって一般化固有値問題よりもむしろ大規模問題の近似解を容易に求めることができる。

(3) 非線形グラフ埋め込み手法である PMvGE において、リンクに外れ値がある場合でもその影響を軽減するロバスト手法を提案した (Okuno and Shimodaira, 2019; AISTATS)。ロバスト化の程度 (パラメータ  $\beta$ ) を増やすと、外れ値の影響が軽減する。右図では、通常の場合 (最尤推定  $\beta=0$ ) に比べて、ロバスト化 ( $\beta=1$ ) すると 3 つのクラスが分離される様子が見える。さらにこれを一般化して、最尤推定における Kullback-Leiber ダイバージェンスを Bregman ダイバージェンスに置き換える手法を提案 (Okuno and Shimodaira, 2020; Neural Networks) した。



(4) グラフにおけるリンクは 2 個のノード間の関連性をあらわすが、これを一般化したハイパーリンクは 3 個以上のノード間の関連性をあらわす。そこで PMvGE によるグラフ埋め込みをハイパーグラフの埋め込みへ一般化する手法 Hyperlink Regression を提案した (Okuno and Shimodaira, 2020; Neural Networks)。

(5) 共通空間に埋め込んだベクトル間の類似度をどのように測るのかについて検討して、PMvGE のようなモデルで表現できる類似度関数のクラスを明らかにした。当初の PMvGE ではベクトル



間の内積を用いていた. このように限定された類似度関数 (内積類似度 IPS) であっても, 十分に大きなニューラルネットワークを用いることによって任意の正定値カーネルを表現できることを証明した (Okuno et al., 2018; ICML). これにバイアス項を加えるモデル (シフト付き内積類似度 SIPS) は任意の条件付き正定値カーネルを表現できる (Okuno et al., 2019; AISTATS). さらに2つの IPS モデルの差 (IPDS) によって任意の不定値カーネルが表現できること (Okuno et al., 2019; AISTATS) を示した. IPDS を効率よく学習するには, 負の値を重みに許す重み付き内積モデル (WIPS) を用いるのが良い (Kim et al., 2019; IJCAI).

## 5. 主な発表論文等

〔雑誌論文〕 計15件（うち査読付論文 15件 / うち国際共著 0件 / うちオープンアクセス 10件）

|  |                         |
|--|-------------------------|
| 1. 著者名<br>Geewook Kim, Kazuki Fukui, Hidetoshi Shimodaira  | 4. 巻<br>-               |
| 2. 論文標題<br>Segmentation-free Compositional n-gram Embedding  | 5. 発行年<br>2019年         |
| 3. 雑誌名<br>Proceedings of 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT) | 6. 最初と最後の頁<br>3207-3215 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>10.18653/v1/N19-1324  | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である）  | 国際共著<br>-               |
| 1. 著者名<br>Geewook Kim, Akifumi Okuno, Kazuki Fukui, Hidetoshi Shimodaira   | 4. 巻<br>-               |
| 2. 論文標題<br>Representation Learning with Weighted Inner Product for Universal Approximation of General Similarities                         | 5. 発行年<br>2019年         |
| 3. 雑誌名<br>Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19                             | 6. 最初と最後の頁<br>5031-5038 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>10.24963/ijcai.2019/699   | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である）  | 国際共著<br>-               |
| 1. 著者名<br>Thong Pham and Paul Sheridan and Hidetoshi Shimodaira}   | 4. 巻<br>92              |
| 2. 論文標題<br>PAFit: An R Package for the Non-Parametric Estimation of Preferential Attachment and Node Fitness in Temporal Complex Networks  | 5. 発行年<br>2020年         |
| 3. 雑誌名<br>Journal of Statistical Software  | 6. 最初と最後の頁<br>-         |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>10.18637/jss.v092.i03   | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である）  | 国際共著<br>-               |
| 1. 著者名<br>Akifumi Okuno and Hidetoshi Shimodaira   | 4. 巻<br>126             |
| 2. 論文標題<br>Hyperlink regression via Bregman divergence   | 5. 発行年<br>2020年         |
| 3. 雑誌名<br>Neural Networks  | 6. 最初と最後の頁<br>362 - 383 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>10.1016/j.neunet.2020.03.026  | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難   | 国際共著<br>-               |

|   |                              |
|---|------------------------------|
| 1. 著者名<br>Masaaki Inoue, Thong Pham, Hidetoshi Shimodaira   | 4. 巻<br>-                    |
| 2. 論文標題<br>Joint Estimation of the Non-parametric Transitivity and Preferential Attachment Functions in Scientific Co-authorship Networks | 5. 発行年<br>2020年              |
| 3. 雑誌名<br>Journal of Informetrics   | 6. 最初と最後の頁<br>Article 101042 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>10.1016/j.joi.2020.101042   | 査読の有無<br>有                   |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)  | 国際共著<br>-                    |

|   |                       |
|---|-----------------------|
| 1. 著者名<br>Hidetoshi Shimodaira, Haruyoshi Maeda   | 4. 巻<br>70            |
| 2. 論文標題<br>An information criterion for model selection with missing data via complete-data divergence                        | 5. 発行年<br>2018年       |
| 3. 雑誌名<br>Annals of the Institute of Statistical Mathematics  | 6. 最初と最後の頁<br>421-438 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br><a href="https://doi.org/10.1007/s10463-016-0592-7">https://doi.org/10.1007/s10463-016-0592-7</a> | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難  | 国際共著<br>-             |

|   |                         |
|---|-------------------------|
| 1. 著者名<br>Akifumi Okuno, Tetsuya Hada, Hidetoshi Shimodaira   | 4. 巻<br>80              |
| 2. 論文標題<br>A probabilistic framework for multi-view feature learning with many-to-many associations via neural networks | 5. 発行年<br>2018年         |
| 3. 雑誌名<br>Proceedings of the 35th International Conference on Machine Learning, PMLR                                    | 6. 最初と最後の頁<br>3888-3897 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>なし  | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)  | 国際共著<br>-               |

|   |                       |
|---|-----------------------|
| 1. 著者名<br>Masaaki Inoue, Thong Pham, Hidetoshi Shimodaira   | 4. 巻<br>IX            |
| 2. 論文標題<br>Transitivity vs Preferential Attachment: Determining the Driving Force Behind the Evolution of Scientific Co-Authorship Networks | 5. 発行年<br>2018年       |
| 3. 雑誌名<br>International Conference on Complex Systems, ICCS 2018: Unifying Themes in Complex Systems  | 6. 最初と最後の頁<br>262-271 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br><a href="https://doi.org/10.1007/978-3-319-96661-8_28">https://doi.org/10.1007/978-3-319-96661-8_28</a>         | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難  | 国際共著<br>-             |

|   |                   |
|---|-------------------|
| 1. 著者名<br>Shinpei Imori, Hidetoshi Shimodaira   | 4. 巻<br>21        |
| 2. 論文標題<br>An Information Criterion for Auxiliary Variable Selection in Incomplete Data Analysis              | 5. 発行年<br>2019年   |
| 3. 雑誌名<br>Entropy   | 6. 最初と最後の頁<br>281 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br><a href="https://doi.org/10.3390/e21030281">https://doi.org/10.3390/e21030281</a> | 査読の有無<br>有        |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)  | 国際共著<br>-         |

|  |                       |
|--|-----------------------|
| 1. 著者名<br>Akifumi Okuno, Geewook Kim, Hidetoshi Shimodaira   | 4. 巻<br>89            |
| 2. 論文標題<br>Graph Embedding with Shifted Inner Product Similarity and Its Improved Approximation Capability | 5. 発行年<br>2019年       |
| 3. 雑誌名<br>Proceedings of Machine Learning Research, PMLR   | 6. 最初と最後の頁<br>644-653 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>なし   | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)   | 国際共著<br>-             |

|   |                       |
|---|-----------------------|
| 1. 著者名<br>Akifumi Okuno, Hidetoshi Shimodaira             | 4. 巻<br>89            |
| 2. 論文標題<br>Robust Graph Embedding with Noisy Link Weights | 5. 発行年<br>2019年       |
| 3. 雑誌名<br>Proceedings of Machine Learning Research, PMLR  | 6. 最初と最後の頁<br>664-673 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>なし                            | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)                    | 国際共著<br>-             |

|   |                       |
|---|-----------------------|
| 1. 著者名<br>Geewook Kim, Kazuki Fukui and Hidetoshi Shimodaira  | 4. 巻<br>4             |
| 2. 論文標題<br>Word-like character n-gram embedding   | 5. 発行年<br>2018年       |
| 3. 雑誌名<br>Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-generated Text | 6. 最初と最後の頁<br>148-152 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>なし  | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)  | 国際共著<br>-             |

|  |                       |
|--|-----------------------|
| 1. 著者名<br>Patrick Bloebaum, Dominik Janzing, Takashi Washio, Shohei Shimizu, Bernhard Schoelkopf                   | 4. 巻<br>84            |
| 2. 論文標題<br>A novel principle for causal inference in data with small error variance                                | 5. 発行年<br>2018年       |
| 3. 雑誌名<br>Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics, PMLR | 6. 最初と最後の頁<br>900-909 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>なし   | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスとしている (また、その予定である)   | 国際共著<br>-             |

|   |                         |
|---|-------------------------|
| 1. 著者名<br>Shimizu Shohei  | 4. 巻<br>20              |
| 2. 論文標題<br>Non-Gaussian Methods for Causal Structure Learning   | 5. 発行年<br>2018年         |
| 3. 雑誌名<br>Prevention Science  | 6. 最初と最後の頁<br>431 ~ 441 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br><a href="https://doi.org/10.1007/s11121-018-0901-x">https://doi.org/10.1007/s11121-018-0901-x</a> | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難  | 国際共著<br>-               |

|  |                       |
|--|-----------------------|
| 1. 著者名<br>福井一輝, 奥野彰文, 下平英寿                                 | 4. 巻<br>D99           |
| 2. 論文標題<br>マッチング相関分析を用いた画像-マルチタグ間の相互検索                     | 5. 発行年<br>2016年       |
| 3. 雑誌名<br>電子情報通信学会論文誌                                      | 6. 最初と最後の頁<br>774-777 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>10.14923/transinfj.2015IUL0005 | 査読の有無<br>有            |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難                     | 国際共著<br>-             |

〔学会発表〕 計48件 (うち招待講演 2件 / うち国際学会 17件)

|  |
|--|
| 1. 発表者名<br>田中卓磨; 奥野彰文; 福井一輝; Kim Geewook; 下平英寿 |
| 2. 発表標題<br>マルチスケールk-近傍法を用いた画像のタグ推定             |
| 3. 学会等名<br>第22回情報論的学習理論ワークショップ                 |
| 4. 発表年<br>2019年                                |



|  |
|--|
| 1. 発表者名<br>水谷守裕; 奥野彰文; 福井一輝; Kim Geewook; 金沢朋実; 白石 友一; 岡田眞里子; 下平英寿 |
| 2. 発表標題<br>グラフと近傍グラフの確率的同時埋め込みによるマルチモーダルデータの可視化                    |
| 3. 学会等名<br>第22回情報論的学習理論ワークショップ                                     |
| 4. 発表年<br>2019年  |

|  |
|--|
| 1. 発表者名<br>奥野 彰文, Geewook Kim, 下平 英寿         |
| 2. 発表標題<br>ニューラルネットワークを用いたグラフ埋め込みの表現 能力とその拡張 |
| 3. 学会等名<br>統計関連学会連合大会                        |
| 4. 発表年<br>2019年                              |

|  |
|--|
| 1. 発表者名<br>奥野 彰文, 下平 英寿                                |
| 2. 発表標題<br>Hyperlink Regression via Bregman Divergence |
| 3. 学会等名<br>統計関連学会連合大会                                  |
| 4. 発表年<br>2019年  |

|  |
|--|
| 1. 発表者名<br>Hidetoshi Shimodaira  |
| 2. 発表標題<br>Multiscale Bootstrap for Selective Inference with Applications to Model Selection |
| 3. 学会等名<br>Data Science, Statistics and Visualization (DSSV2019) (招待講演) (国際学会)               |
| 4. 発表年<br>2019年  |

|  |
|--|
| 1. 発表者名<br>Hidetoshi Shimodaira  |
| 2. 発表標題<br>Multiview Graph Embedding as a Generalization of Canonical Correlation Analysis |
| 3. 学会等名<br>62nd ISI World Statistics Congress 2019 (ISI WSC 2019) (招待講演) (国際学会)            |
| 4. 発表年<br>2019年  |

|   |
|---|
| 1. 発表者名<br>Masaaki Inoue, Thong Pham, Hidetoshi Shimodaira  |
| 2. 発表標題<br>Statistical Estimation of the Effects of First and Second Order Local Structures on Growth of Complex Networks |
| 3. 学会等名<br>ACML 2019 Workshop on Statistics & Machine Learning Researchers in Japan (国際学会)                                |
| 4. 発表年<br>2019年   |

|                                    |
|------------------------------------|
| 1. 発表者名<br>Kim Geewook, 横井祥, 下平英寿  |
| 2. 発表標題<br>単語埋め込みの二種類の加法構成性        |
| 3. 学会等名<br>言語処理学会第26回年次大会(NLP2020) |
| 4. 発表年<br>2020年                    |

|  |
|--|
| 1. 発表者名<br>Akifumi Okuno and Hidetoshi Shimodaira  |
| 2. 発表標題<br>On representation power of neural network-based graph embedding and beyond                  |
| 3. 学会等名<br>ICML2018 workshop Theoretical Foundations and Applications of Deep Generative Models (国際学会) |
| 4. 発表年<br>2018年  |

|                                      |
|--------------------------------------|
| 1. 発表者名<br>Geewook Kim, 奥野 彰文, 下平 英寿 |
| 2. 発表標題<br>擬ユークリッド空間への単語埋め込み         |
| 3. 学会等名<br>言語処理学会 第25回年次大会           |
| 4. 発表年<br>2019年                      |

|   |
|---|
| 1. 発表者名<br>奥野 彰文, 下平英寿                  |
| 2. 発表標題<br>ニューラルネットワークを用いた異種データのグラフ埋め込み |
| 3. 学会等名<br>2018年度 統計関連学会連合大会            |
| 4. 発表年<br>2018年                         |

|  |
|--|
| 1. 発表者名<br>寺田吉彦, 下平英寿                                    |
| 2. 発表標題<br>マルチスケール・ブートストラップによるモデル選択後のselective inference |
| 3. 学会等名<br>2018年度 統計関連学会連合大会                             |
| 4. 発表年<br>2018年  |

|                                      |
|--------------------------------------|
| 1. 発表者名<br>前田篤刀, 伊森晋平, 下平英寿          |
| 2. 発表標題<br>回帰モデルにおける補助変数を活用した推定精度の向上 |
| 3. 学会等名<br>2018年度 統計関連学会連合大会         |
| 4. 発表年<br>2018年                      |

|  |
|--|
| 1. 発表者名<br>井上雅章, Pham THONG, 下平英寿                      |
| 2. 発表標題<br>複雑ネットワーク成長メカニズムにおける優先的選択性と推移性のノンパラメトリック同時推定 |
| 3. 学会等名<br>2018年度 統計関連学会連合大会                           |
| 4. 発表年<br>2018年  |

|                                   |
|-----------------------------------|
| 1. 発表者名<br>福井一輝, 下平英寿             |
| 2. 発表標題<br>単語ベクトルを利用した文書の教師なし行列表現 |
| 3. 学会等名<br>情報論的学習理論と機械学習ワークショップ   |
| 4. 発表年<br>2018年                   |

|                                 |
|---------------------------------|
| 1. 発表者名<br>奥野彰文, 下平英寿           |
| 2. 発表標題<br>ロバストなグラフ埋め込み         |
| 3. 学会等名<br>情報論的学習理論と機械学習ワークショップ |
| 4. 発表年<br>2018年                 |

|   |
|---|
| 1. 発表者名<br>奥野彰文, Geewook Kim, 下平英寿          |
| 2. 発表標題<br>ニューラルネットワークを用いたグラフ埋め込みの表現定理とその拡張 |
| 3. 学会等名<br>情報論的学習理論と機械学習ワークショップ             |
| 4. 発表年<br>2018年                             |

|                                 |
|---------------------------------|
| 1. 発表者名<br>田中卓磨, 奥野彰文, 下平英寿     |
| 2. 発表標題<br>グラフ埋め込みの次数補正とその応用    |
| 3. 学会等名<br>情報論的学習理論と機械学習ワークショップ |
| 4. 発表年<br>2018年                 |

|                          |
|--------------------------|
| 1. 発表者名<br>清水昌平          |
| 2. 発表標題<br>因果探索、予測、そして制御 |
| 3. 学会等名<br>応用統計学会        |
| 4. 発表年<br>2018年          |

|   |
|---|
| 1. 発表者名<br>S. Shimizu   |
| 2. 発表標題<br>Causal discovery, prediction mechanisms, and control   |
| 3. 学会等名<br>The 5th meeting of the Institute of Mathematical Statistics (IMS) meeting series, the IMS Asia Pacific Rim Meeting (IMS-APRM) (国際学会) |
| 4. 発表年<br>2018年   |

|   |
|---|
| 1. 発表者名<br>S. Shimizu   |
| 2. 発表標題<br>Causal discovery, prediction, and control                    |
| 3. 学会等名<br>Causal Modeling and Machine Learning (CaMaL) Workshop (国際学会) |
| 4. 発表年<br>2018年   |

|  |
|--|
| 1. 発表者名<br>Akifumi Okuno and Hidetoshi Shimodaira  |
| 2. 発表標題<br>Leveraging local data structure for multi-view analysis with many-to-many associations    |
| 3. 学会等名<br>Conference of the International Federation of Classification Societies (IFCS-2017) (国際学会) |
| 4. 発表年<br>2017年  |

|  |
|--|
| 1. 発表者名<br>Akifumi Okuno and Hidetoshi Shimodaira  |
| 2. 発表標題<br>Statistical consistency of multi-view correlation analysis with many-to-many associations |
| 3. 学会等名<br>Joint Statistical Meeting 2017 (国際学会)   |
| 4. 発表年<br>2017年  |

|  |
|--|
| 1. 発表者名<br>Akifumi Okuno and Hidetoshi Shimodaira                    |
| 2. 発表標題<br>Robust Multi-view Graph Embedding                         |
| 3. 学会等名<br>International Conference on Robust Statistics 2017 (国際学会) |
| 4. 発表年<br>2017年  |

|   |
|---|
| 1. 発表者名<br>Takamasa Oshikiri  |
| 2. 発表標題<br>Segmentation-Free Word Embedding for Unsegmented Languages                         |
| 3. 学会等名<br>Conference on Empirical Methods in Natural Language Processing (EMNLP 2017) (国際学会) |
| 4. 発表年<br>2017年   |

|   |
|---|
| 1. 発表者名<br>Kazuki Fukui, Takamasa Oshikiri, Hidetoshi Shimodaira                                  |
| 2. 発表標題<br>Spectral Graph-Based Method of Multimodal Word Embedding                               |
| 3. 学会等名<br>Workshop on Graph-based Methods for Natural Language Processing (TextGraphs-11) (国際学会) |
| 4. 発表年<br>2017年   |

|  |
|--|
| 1. 発表者名<br>Tetsuya Hada, Akifumi Okuno, Hidetoshi Shimodaira                             |
| 2. 発表標題<br>Deep Multi-view Representation Learning Based on Adaptive Weighted Similarity |
| 3. 学会等名<br>First International Workshop on Symbolic-Neural Learning (SNL-2017) (国際学会)    |
| 4. 発表年<br>2017年  |

|   |
|---|
| 1. 発表者名<br>Takahiro Ido, Shinpei Imori, Hidetoshi Shimodaira                                      |
| 2. 発表標題<br>An Information Criterion For Prediction With Auxiliary Variables Under Covariate Shift |
| 3. 学会等名<br>IASC-ARS/NZSA 2017 (国際学会)  |
| 4. 発表年<br>2017年   |

|  |
|--|
| 1. 発表者名<br>Akifumi Okuno, Hidetoshi Shimodaira   |
| 2. 発表標題<br>Cross-view link prediction with attribute vectors and its information criterion |
| 3. 学会等名<br>2017年度統計関連学会連合大会  |
| 4. 発表年<br>2017年  |

|   |
|---|
| 1. 発表者名<br>福井 一輝, 押切 孝将, 下平 英寿                |
| 2. 発表標題<br>クロスドメインマッチング相関分析による画像情報を反映した単語埋め込み |
| 3. 学会等名<br>2017年度統計関連学会連合大会                   |
| 4. 発表年<br>2017年                               |

|   |
|---|
| 1. 発表者名<br>寺田 吉彦, 下平 英寿                                 |
| 2. 発表標題<br>マルチスケールブートストラップによる近似的に不偏なselective inference |
| 3. 学会等名<br>2017年度統計関連学会連合大会                             |
| 4. 発表年<br>2017年   |

|                                       |
|---------------------------------------|
| 1. 発表者名<br>井戸 貴大, 伊森 晋平, 下平 英寿        |
| 2. 発表標題<br>共変量シフトにおける補助変数を用いた予測と情報量規準 |
| 3. 学会等名<br>2017年度統計関連学会連合大会           |
| 4. 発表年<br>2017年                       |

|  |
|--|
| 1. 発表者名<br>奥野彰文, 羽田哲也, 下平英寿                        |
| 2. 発表標題<br>属性ベクトルとニューラルネットワークを用いた異種データ間のリンク構造の最尤推定 |
| 3. 学会等名<br>第20回情報論的学習理論ワークショップ                     |
| 4. 発表年<br>2017年                                    |



|   |
|---|
| 1. 発表者名<br>福井一輝, 下平英寿                           |
| 2. 発表標題<br>クロスドメインマッチング相関分析を用いた画像とキャプションの同時埋め込み |
| 3. 学会等名<br>第20回情報論的学習理論ワークショップ                  |
| 4. 発表年<br>2017年                                 |

|   |
|---|
| 1. 発表者名<br>Geewook Kim, 福井 一輝, 羽田 哲也, 下平 英寿 |
| 2. 発表標題<br>単語辞書を併用した単語分割しない単語埋め込み           |
| 3. 学会等名<br>第20回情報論的学習理論ワークショップ              |
| 4. 発表年<br>2017年                             |

|  |
|--|
| 1. 発表者名<br>Kim Geewook, 福井一輝, 羽田哲也, 下平英寿 |
| 2. 発表標題<br>単語らしい文字n-gramの埋め込みによる単語の分散表現  |
| 3. 学会等名<br>言語処理学会第24回年次大会(NLP2018)       |
| 4. 発表年<br>2018年                          |

|                                     |
|-------------------------------------|
| 1. 発表者名<br>小嶋啓右, 廣瀬慧, 下平英寿          |
| 2. 発表標題<br>多ドメインマッチング相関分析のL1正則化法    |
| 3. 学会等名<br>情報論的学習理論と機械学習(IBISML)研究会 |
| 4. 発表年<br>2016年                     |

|  |
|--|
| 1. 発表者名<br>Oshikiri, Takamasa and Fukui, Kazuki and Shimodaira, Hidetoshi              |
| 2. 発表標題<br>Cross-Lingual Word Representations via Spectral Graph Embeddings            |
| 3. 学会等名<br>54th Annual Meeting of the Association for Computational Linguistics (国際学会) |
| 4. 発表年<br>2016年  |

|  |
|--|
| 1. 発表者名<br>Fukui, Kazuki and Okuno, Akifumi and Shimodaira, Hidetoshi                                |
| 2. 発表標題<br>Image and tag retrieval by leveraging image-group links with multi-domain graph embedding |
| 3. 学会等名<br>2016 IEEE International Conference on Image Processing (国際学会)                             |
| 4. 発表年<br>2016年  |

|   |
|---|
| 1. 発表者名<br>下平英寿                               |
| 2. 発表標題<br>データベクトルのマッチングに関する情報統合の多変量解析とその最尤推定 |
| 3. 学会等名<br>統計関連学会連合大会                         |
| 4. 発表年<br>2016年                               |

|                                   |
|-----------------------------------|
| 1. 発表者名<br>押切孝将, 福井一輝, 下平英寿       |
| 2. 発表標題<br>マッチング相関分析を用いた多言語単語埋め込み |
| 3. 学会等名<br>統計関連学会連合大会             |
| 4. 発表年<br>2016年                   |

|   |
|---|
| 1. 発表者名<br>羽田哲也, 下平英寿                       |
| 2. 発表標題<br>多ドメインマッチング相関分析の深層ニューラルネットによる非線形化 |
| 3. 学会等名<br>統計関連学会連合大会                       |
| 4. 発表年<br>2016年                             |

|  |
|--|
| 1. 発表者名<br>奥野彰文, 下平英寿                  |
| 2. 発表標題<br>多対多対応を利用したマッチング相関分析の一致性について |
| 3. 学会等名<br>統計関連学会連合大会                  |
| 4. 発表年<br>2016年                        |

|                                       |
|---------------------------------------|
| 1. 発表者名<br>羽田哲也, 福井一輝, 下平英寿           |
| 2. 発表標題<br>深層クロスドメインマッチング相関分析の提案とその応用 |
| 3. 学会等名<br>第19回情報論的学習理論ワークショップ        |
| 4. 発表年<br>2016年                       |

|                             |
|-----------------------------|
| 1. 発表者名<br>押切孝将, 下平英寿       |
| 2. 発表標題<br>単語分割を経由しない単語埋め込み |
| 3. 学会等名<br>言語処理学会第23回年次大会   |
| 4. 発表年<br>2017年             |

|                                 |
|---------------------------------|
| 1. 発表者名<br>奥野彰文, 下平英寿           |
| 2. 発表標題<br>多対多対応を扱う相関分析の一致性について |
| 3. 学会等名<br>第11回日本統計学会春季集会       |
| 4. 発表年<br>2017年                 |

|  |
|--|
| 1. 発表者名<br>Shimizu, Shohei   |
| 2. 発表標題<br>A non-Gaussian model for causal discovery in the presence of hidden common causes |
| 3. 学会等名<br>Munich Workshop on Causal Inference and Information Theory (国際学会)                 |
| 4. 発表年<br>2016年  |

|  |
|--|
| 1. 発表者名<br>Shimizu, Shohei   |
| 2. 発表標題<br>A non-Gaussian approach for causal structure learning in the presence of hidden common causes |
| 3. 学会等名<br>CRM Workshop: Statistical Causal Inference and its Applications to Genetics (国際学会)            |
| 4. 発表年<br>2016年  |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

|   |
|---|
| Shimodaira Lab<br><a href="http://stat.sys.i.kyoto-u.ac.jp/">http://stat.sys.i.kyoto-u.ac.jp/</a> |
|---|

## 6. 研究組織

|                   | 氏名<br>(ローマ字氏名)<br>(研究者番号)                       | 所属研究機関・部局・職<br>(機関番号)                         | 備考 |
|-------------------|---|---|----|
| 研究<br>分<br>担<br>者 | 清水 昌平<br><br>(Shimizu Shohei)<br><br>(10509871) | 滋賀大学・データサイエンス学部・教授<br><br><br><br><br>(14201) |    |