

ハイパーパラメータチューニングにおける 不確実性の影響削減のための再評価戦略

今村 秀明 (東大) 渡部 正樹 (Preferred Networks) 秋葉 拓哉 (Preferred Networks)

目的

同一パラメータの評価結果が試行ごとにバラついてしまう影響で、誤って性能の悪いパラメータが提案されてしまう問題を解決する。

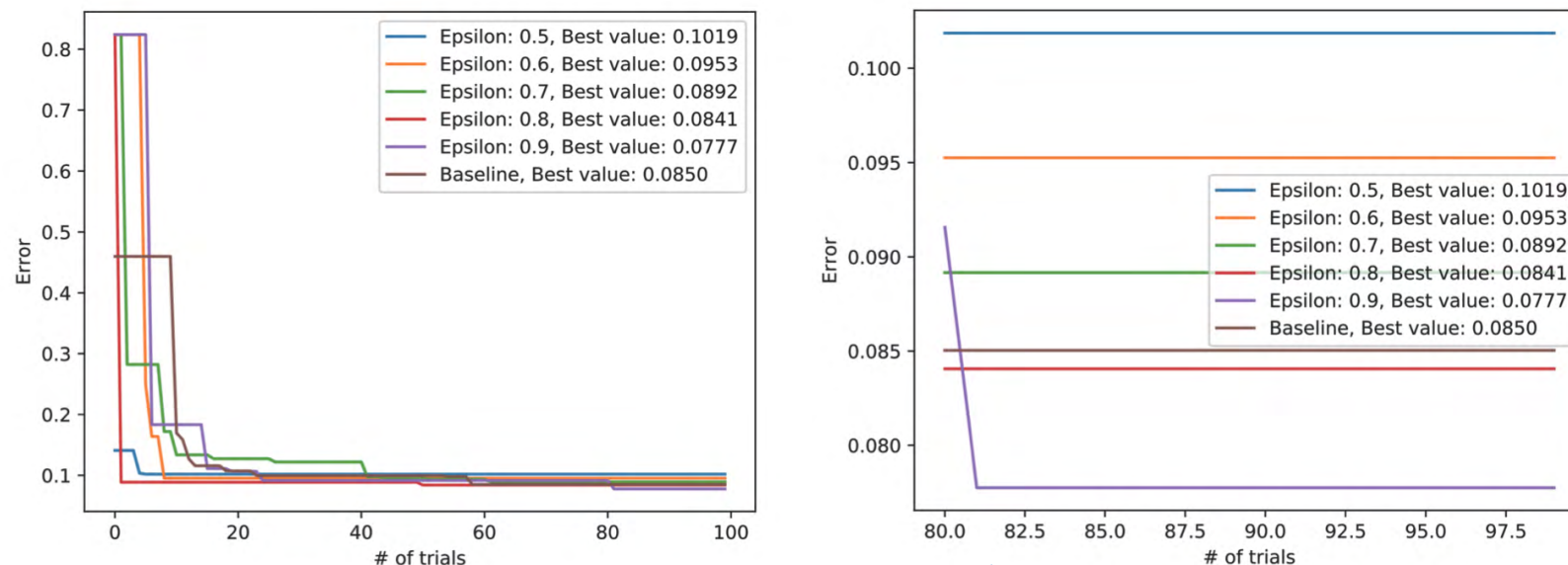
手法

同一のパラメータを確率的に再評価することで評価結果の修正を行う。

結果

幾つかの確率的再評価戦略を提案し性能が高いことを実験的に示した。

再評価しない場合と再評価を一定確率($1 - \epsilon$)でする場合の比較実験



トライアル数が80~100の部分に注目

ロバストなグラフ埋め込み

奥野 彰文 (京大・理研AIP) 下平英寿 (京大・理研AIP)

目的

“隣接行列に含まれるノイズ”に影響を受けにくいグラフ埋め込みを実現する

手法

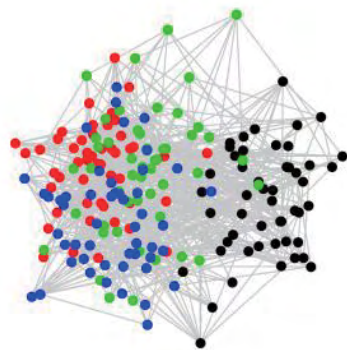
対数尤度の代わりに, 提案する Moment Matching β -Score Function (MMBSF) を最適化

結果

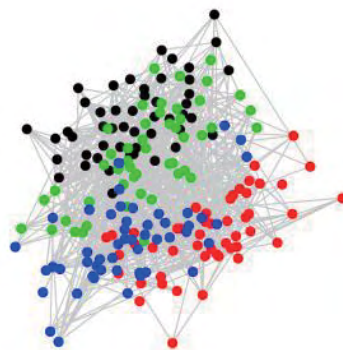
MMBSFは上記ノイズの影響を受けにくく, グラフ埋め込みがロバストになる.

観測値: Link weights $w_{ij} = w_{ji} \in \mathbb{R}_{\geq 0}$, Data vectors $\mathbf{x}_i \in \mathbb{R}^p$ ($i, j = 1, 2, \dots, n$)

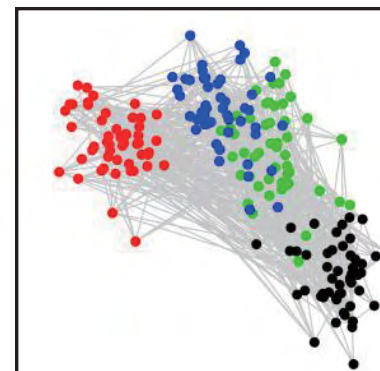
設定: $E(w_{ij} \mid \mathbf{x}_i, \mathbf{x}_j) = \underbrace{\mu_*(\mathbf{x}_i, \mathbf{x}_j)}_{\text{真の期待値}} + \underbrace{\alpha \eta_*(\mathbf{x}_i, \mathbf{x}_j)}_{\text{ノイズ}}$



既存手法
(対数尤度最大化)



提案手法
(MMBSF最大化, $\beta = 0.5$)



提案手法
(MMBSF最大化, $\beta = 1.0$)

提案手法は
ノイズの影響を無視して
潜在的なクラスターを
復元できている

図: 埋め込みを2次元で可視化
提案法は $\beta = 0$ のとき既存手法
に対応し $\beta > 0$ のときロバスト

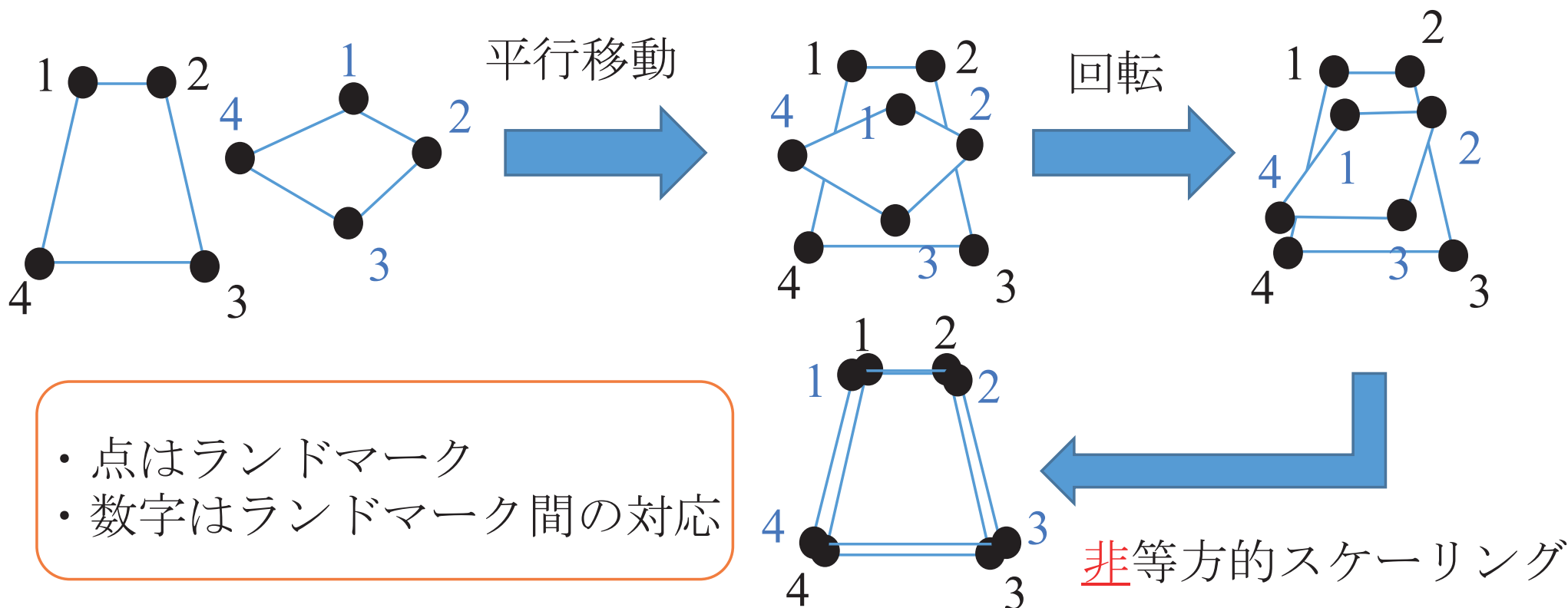
形状検索のための距離計量学習

岡本 司 岩田 一貴 末松 伸朗 三村 和史 (広島市立大)

目的 平行移動, 回転, 非等方的スケーリングを考慮した形状間距離を提案

手法 非等方的スケーリングを距離計量学習によって最適化

結果 形状間距離に基づく形状検索において, 提案する距離は高性能



形状局所記述子を用いた形状整合の新しい枠組み

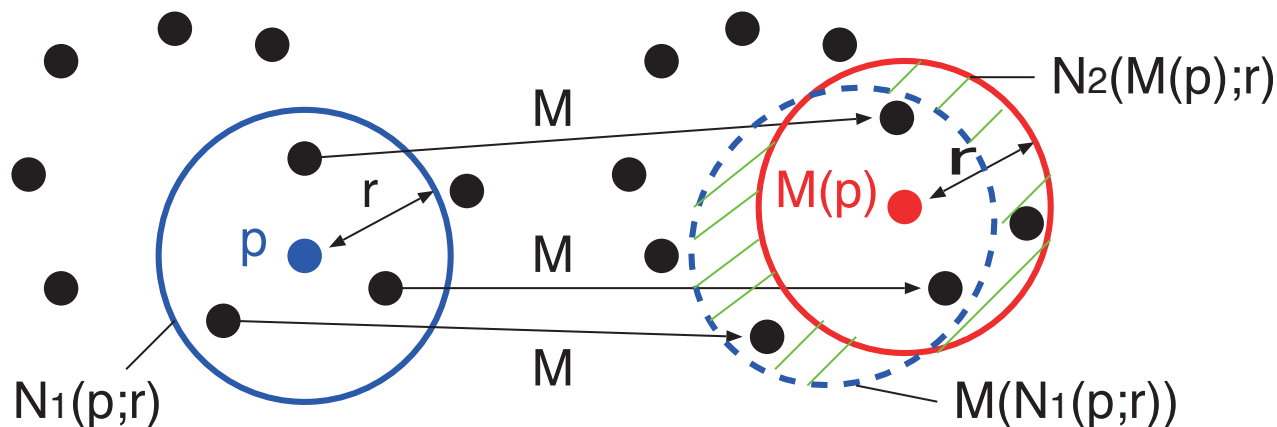
山本 大貴, 岩田 一貴, 末松 伸朗, 三村 和史 (広島市立大学)

目的 従来の局所記述子の識別性能を向上させるような形状整合を提案

手法 近傍標本点の対応先を考慮するような整合を考え, その計算を効率化

結果 計算時間は少し多くかかるが, 局所記述子の性能を向上させることに成功

近傍標本点の対応先を考慮した追加整合コスト



左の形状の標本点 p が右の形状の標本点 $M(p)$ に対応するとき, p の近傍標本点 (青の実線円内の点) の対応先の集合 (青の破線円) と $M(p)$ の近傍標本点の集合 (赤の実線円) の対称差 (緑の斜線部) を考え, その要素数を整合コストに組み入れる.

巨大ランダム行列を用いた構造の変化点検知

伊藤 克哉(東京大学) 木下 僚(PFN) 小松 智希(PFN) 吉川 真史(PFN)

目的 潜在時系列と観測時系列の関係の変化を検知する

手法 時系列データを巨大なランダム行列と疎な行列に分解し、重み行列を推定

結果 人工的に生成したデータで、手法の有効性を確認

[問題設定] t^* : 変化点を検知したい **[手法]** $Y = [Y_i, \dots, Y_j]$ を分解する

• $Y_t = A_t Z_t$: 観測時系列

• Z_t : 潜在時系列

• A_t : 重み行列

• $A_0 = \dots = A_{t^*-1}$

$\neq A_{t^*} = \dots = A_T$

• $A_0 \neq A_T$ だが、平均・分散は同じ

• 例:

$A_{ij} \sim N(0, 1) \rightarrow A_{ij} \sim t_{10}(0, 1)$

• M : スパースな係数行列

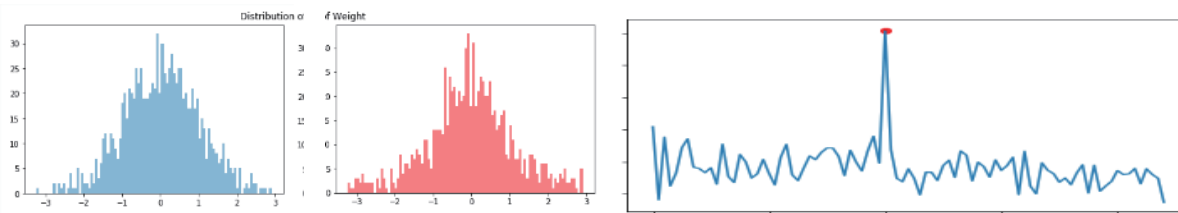
• N : 巨大なランダム行列

• $Y = MN$ を満たすものを探す

1. minimize $\|M\|_{L^1}$ s.t. $\frac{\|Y - MN\|_{L^2}}{\|Y\|_{L^2}} \leq \alpha$

2. $\rightarrow M \simeq A_t, N \simeq Z_t \rightarrow M$ の変化を検知

[結果] 正規分布からt分布への変化等を検知



Graph Convolutionの拡張に関する一検討

石黒 勝彦(PFN) 前田 新一(PFN)

目的

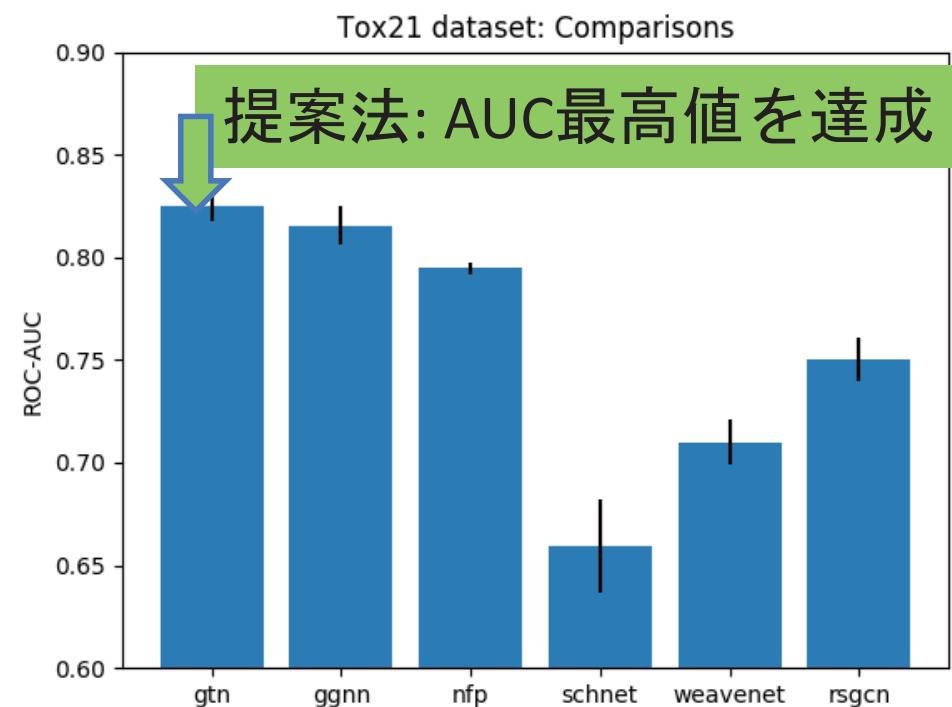
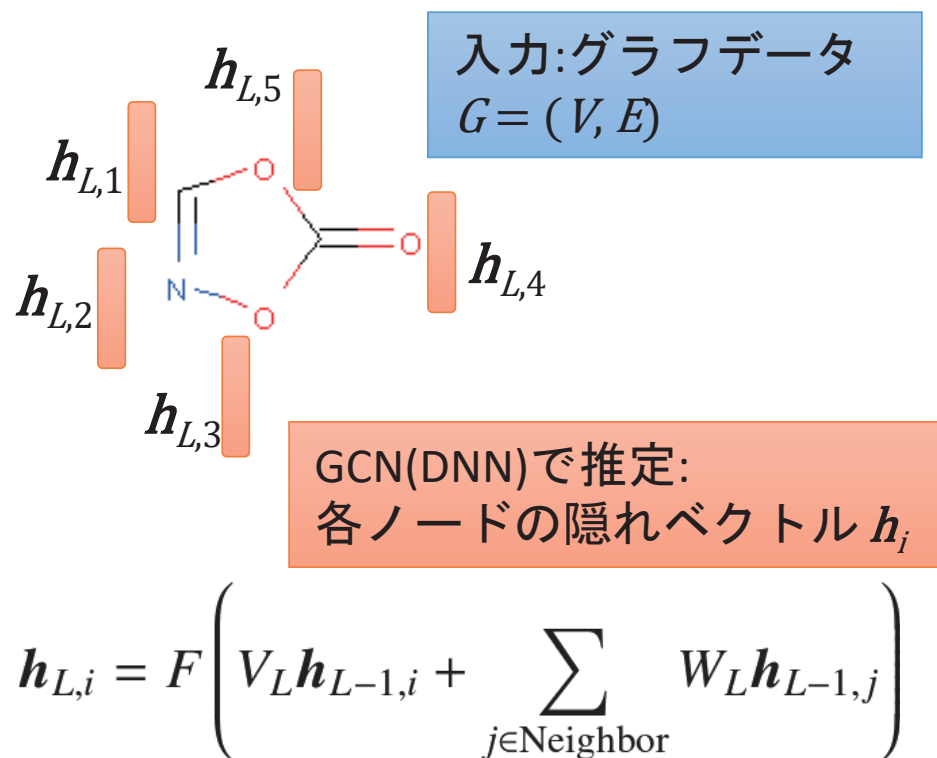
Graph Convolution Networkによるグラフデータ解析の性能向上

手法

Gate/Attention, super node, multi-edge typeといった技術要素を利用

結果

分子グラフデータの性質予測問題においてSOTA達成



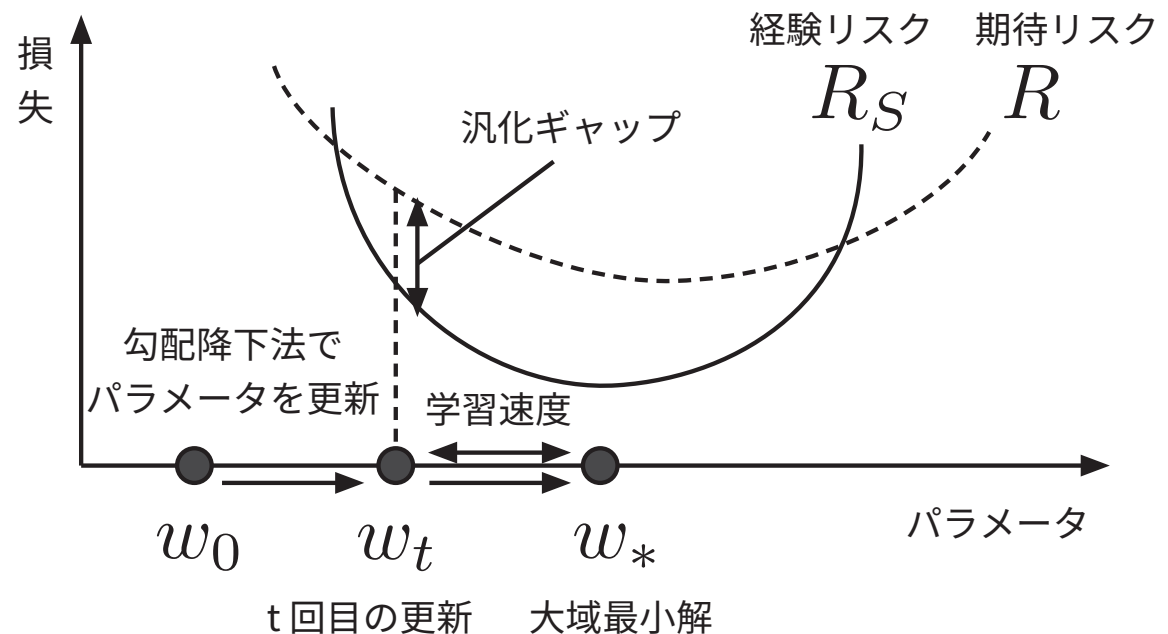
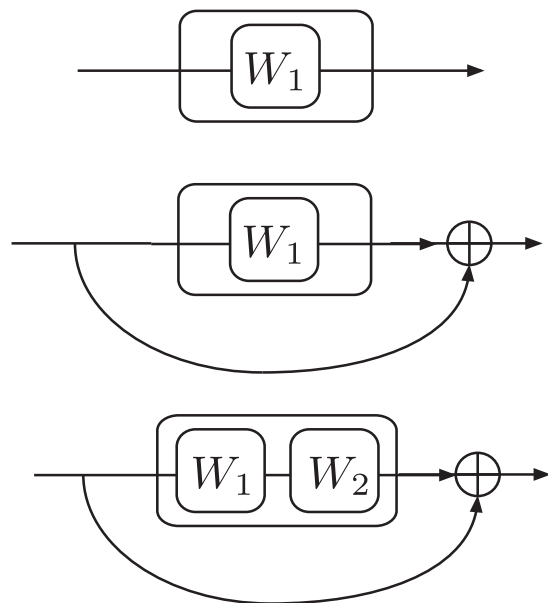
ResNetのショートカットが 学習速度と汎化ギャップに及ぼす影響の理論解析

古庄泰隆 (NAIST) Tongliang Liu (Univ. Sydney) 池田和司 (NAIST)

目的 ResNetのショートカットが学習速度と汎化ギャップへ及ぼす影響を解析

手法 MLPとショートカットが1層、2層スキップするResNetの上記性質を比較

結果 1層スキップするResNetはMLPと変わらない
2層スキップするResNetは学習速度は遅いが、汎化ギャップは小さくなる



乗算ノイズによる勾配降下的高速ロバスト化

Matthew J. Holland (大阪大学)

目的

ERM-GDにまさるデータ分布への頑健性と学習効率保証を両立させる

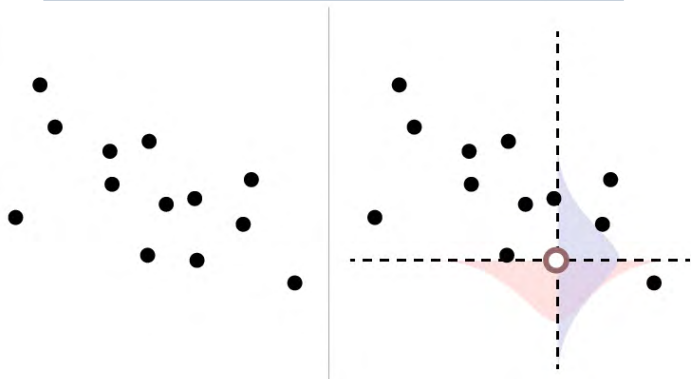
手法

解析的に平滑化した乗算ノイズを加味したリスク勾配推定をGDに組み込む

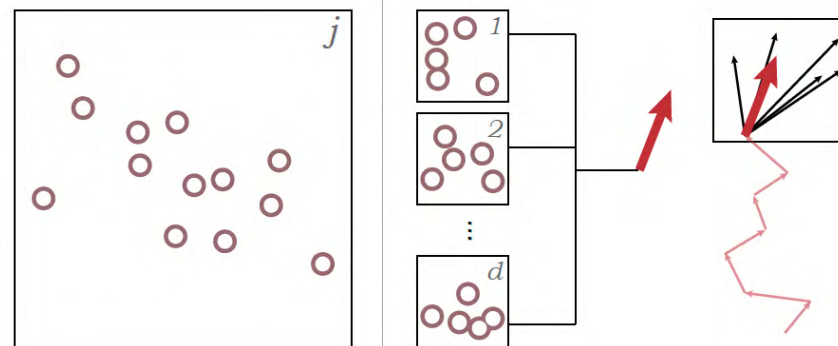
結果

強い性能保証をそのままに、従来のロバストGDと比較にならないほど速い

ノイズをデータに掛けて、
re-scale & truncateしたのちに
真の期待値を取る



数値積分はせず、解析的に
求まる値を次元ごとに算出し、
近似ベクトルを構成していく



ポイント: 頑健化ベクトルの閉形式 + PAC-Bayes不等式

最後は、通常の最急降下の
更新式に代入するだけ

skew-t分布を適用した状態空間モデルによる 将来の健康状態予測

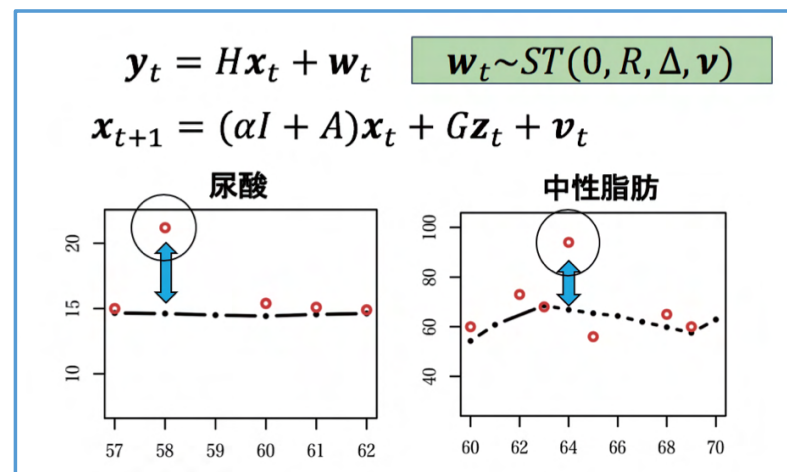
長谷川 嵩矩 (東大医科研) 新井田 厚司 (東大医科研)
山口 類 (東大医科研) 井元 清哉 (東大医科研)

目的 血液検査値特有のノイズにロバストなモデルで健康状態の将来予測を行う

手法 Skew-t分布型の観測ノイズを持つ状態空間モデルを導入する

結果 健診前の食事や運動に敏感な検査値に対してもロバストに予測を行えた

- 状態空間モデルを用いた構造推定は様々な問題に対して行われてきた
- 健康診断データのノイズは偏りがあり裾が重いので拡張が必要
 - 尿酸や中性脂肪など、前日の食事や運動次第で高値が出がちになる (下図参照)
- skew-t分布型の観測ノイズを仮定し、階層表現を用いて同時確率分布を表現
- 変分近似を用いて隠れ変数とノイズの条件付き確率分布を計算
- L1制約付きEMアルゴリズムを用いてパラメータを推定
- 推定したパラメータを用いて健診受診者の将来の健康状態を予測



画像分類における適応的なSGD挙動の解析

今川和樹(茨城大理工) 竹田晃人(茨城大理工)

目的

近年考案されたR-SGDの挙動と画像分類での誤差関数の性質を考察する.

手法

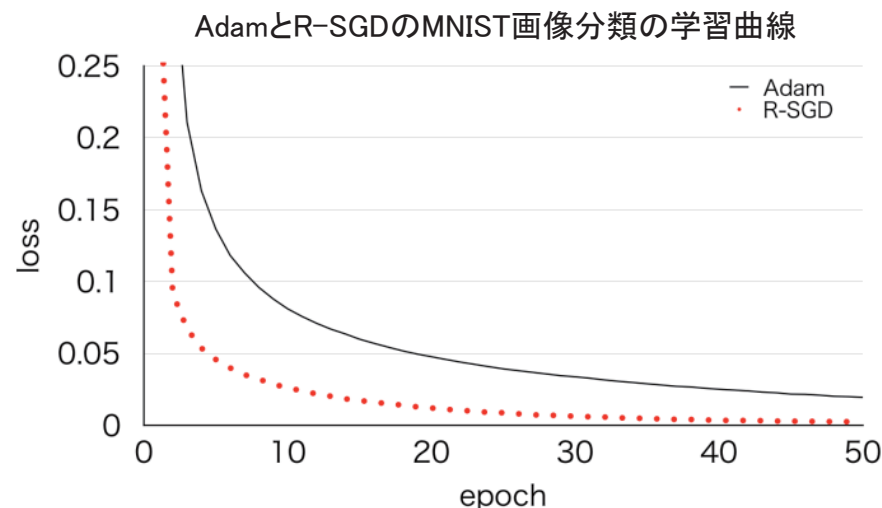
Nudged Elastic Band(NEB)法を用い訓練済みパラメータ間の経路を解析する.

結果

経路間に山が見られR-SGDはそれを超えているような挙動がある？

• Reinforced-SGD

- 基本的にはSGDと同じだが, パラメータ更新において確率的に過去の勾配を利用することにより鞍点による影響を減らす.



- MNIST画像分類
- AdamとR-SGDの比較
- AdamよりもR-SGDの方が収束が早く, lossの値も低い結果が出た



NEBを用いて挙動の解析

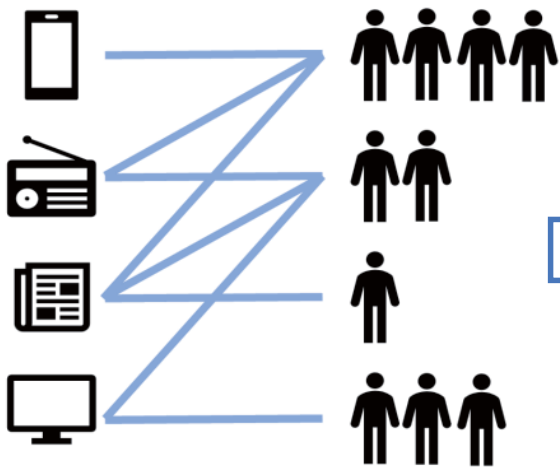
重み付き最大被覆問題への確率伝搬法の適用

喜多野 広貴（茨城大） 竹田 晃人（茨城大）

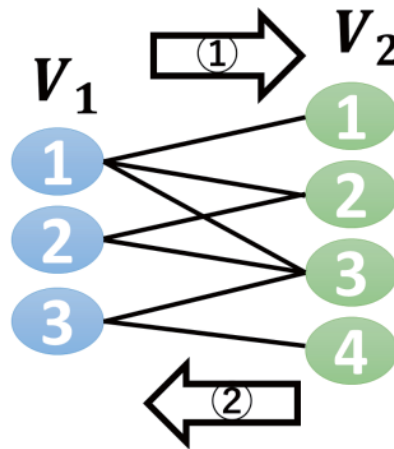
目的 2部グラフ上の重み付き最大被覆問題を解く

手法 確率伝搬法の適用

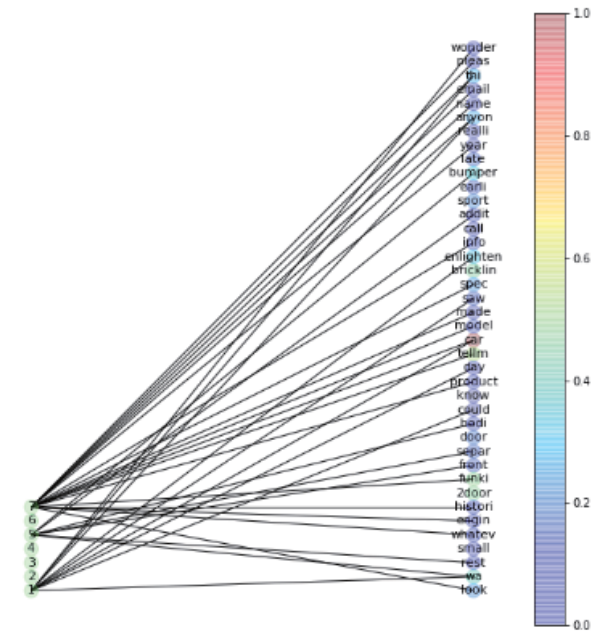
結果 確率伝搬法の適用により精度の向上が見られた



重み付き最大被覆問題



確率伝搬法の適用



文書要約問題への応用

Model Selection with a Renewable Prior Distribution for Intractable Likelihood

Takafumi Kajihara*(NEC), Motonobu Kanagawa*(MPI),
Yuuki Nakaguchi (NEC), Kanishka Khandelwal(NEC), Kenji Fukumizu (ISM)

目的

Model selection with intractable likelihood and unreliable prior knowledge

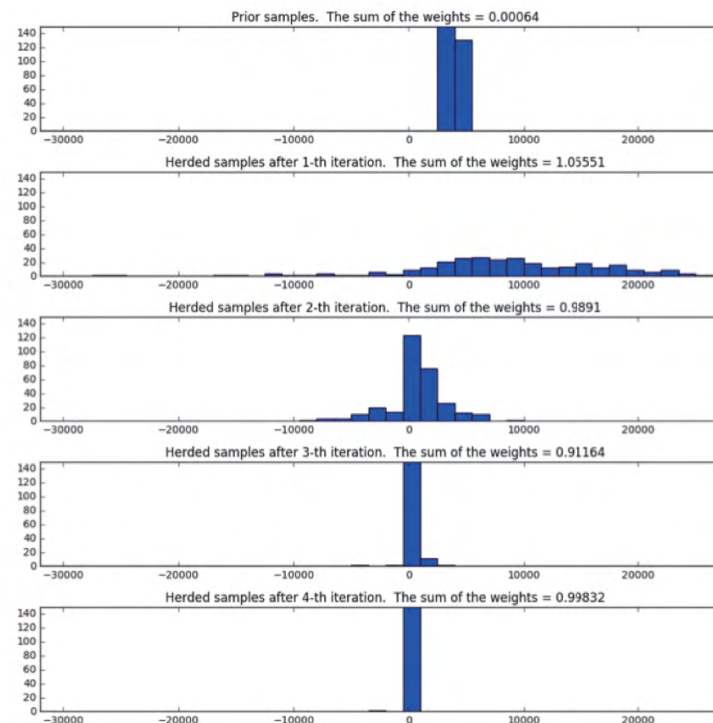
手法

Simultaneous estimation of a model and its parameter

結果

Robust model selection with a misspecified or too-wide parameter prior

Iteration



そのクラスタ信用できますか？ ークラスタ分割に対する妥当性の統計的検証ー

井上 茂乗(名工大) 梅津 佑太(名工大) 竹内 一郎(名工大・理研・NIMS)

目的 K-Means Clusteringのクラスタ分割に対する統計的な評価方法の確立

手法 “各クラスタの母平均が同じ”という仮説に基づく検定を, **Selective Inference** の枠組みで考える

結果 クラスタリングの影響を考慮したexactな検定を構築

ー クラスタの母平均に関する仮説

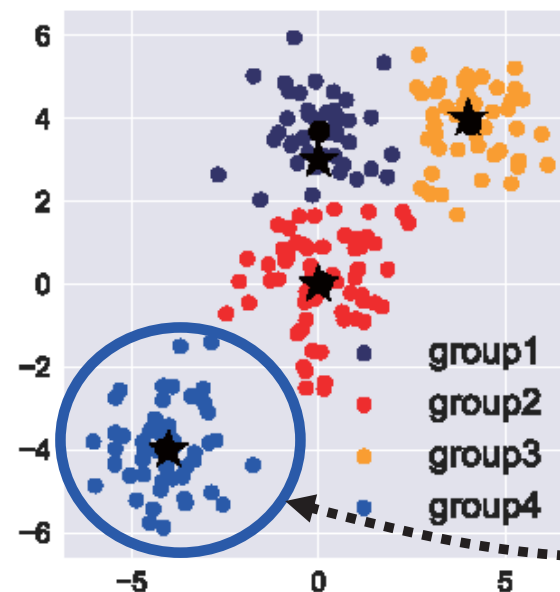
$$H_0^{(a,b)} : \mu_a = \mu_b$$

ー **Point:**

検証すべき仮説がクラスタリング
結果に依存し, **仮説がランダム**

➡ 通常のカイ二乗検定ではダメ

➡ **Selective Inference**



K-Means Clusteringを適用

クラスタ4は他とは母平均
が異なるクラスタといえる

1	0.241	0.066	0.022
2		0.000	0.006
3			0.000

クラスタ同士の比較のp値

データ均一化による回帰性能の向上を目指して

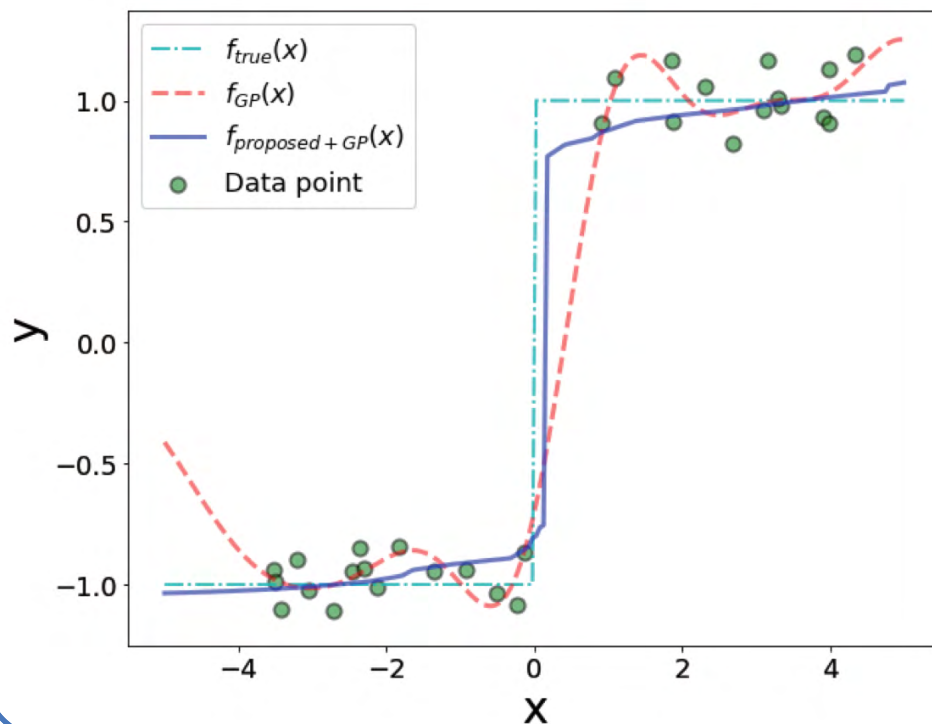
井手 敦也 角所 考 岡留 剛 (関西学院大学)

目的 不連続を含む関数に対する精度の高い回帰

手法 可逆な非線形変換を目標値に施した空間での回帰と, その後の逆変換

結果 既存手法よりも真の関数に近い回帰関数

提案手法と既存手法の比較



➤ 回帰手法: ガウス過程回帰

➤ スコア

手法	MAE	RMSE
GP	0.180	0.345
提案手法+GP	0.082	0.230

➤ 手順

- 1) データ点 y_i を用い, 分布関数 $p(y)$ を推定
- 2) $p(y)$ より累積分布関数 $h(y)$ を算出
- 3) データ点を $\hat{y}_i = h(y_i)$ で変換
- 4) $(x_i, \hat{y}_i)$ を用い既存の回帰手法に適用
- 5) 予測値を $y = h^{-1}(\hat{y})$ によって得る

ベイズ的測度最適化

林 勝悟 (京大) 金川 元信 (マックスプランク研究所) 鹿島 久嗣 (京大・理研AIP)

目的

ブラックボックス関数 f の積分値に関してある測度 $\nu \in \mathcal{M}$ を効率的に最適化するために関数の次の評価点を決定

手法

測度を入力とする獲得関数

結果

ディラック測度を避ける方法論が課題

- 目的関数 : $z^* = \max_{\nu \in \mathcal{M}} \int f(x) d\nu(x)$
- 応用①(ディラック測度) : ベイズ最適化
- 応用②(確率測度) : 入力ノイズありベイズ最適化, Ecological Inference
- 応用③(ルベーク測度) : 最大積分値探索
- 特徴 : $f \sim GP(0, k)$ と仮定すれば $z \sim GP(0, \ell)$, $\ell(\nu, \nu') := \iint k(x, x') d\nu(x) d\nu'(x')$
- 手法 : ベイズ最適化の獲得関数を利用して次に評価を行いたい測度を決定し, その積分値の予測分散を最小化する入力を次の評価点として出力
- 課題 : 自明な解, ディラック測度を避けながら最適化を行いたい

Training Discriminative Model for Anomaly Detection through Generative Adversarial Network

Hiroataka Hachiya (Wakayama Univ.)

目的

Improve recent GAN based abnormal detection methods

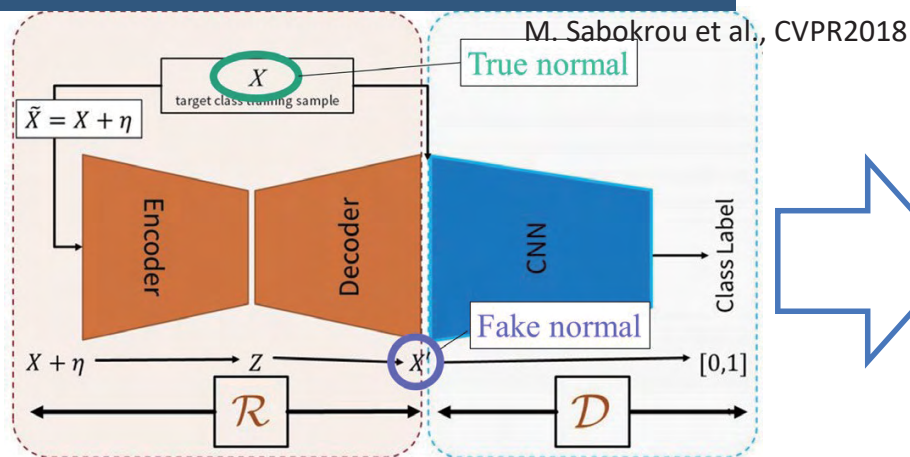
手法

Generate **pseudo abnormal data** using auto-encoder to train anomaly detector

結果

The performance improvement is shown through the experiment with MNIST

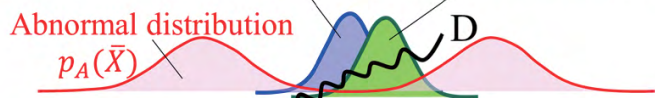
Existing GAN based approach



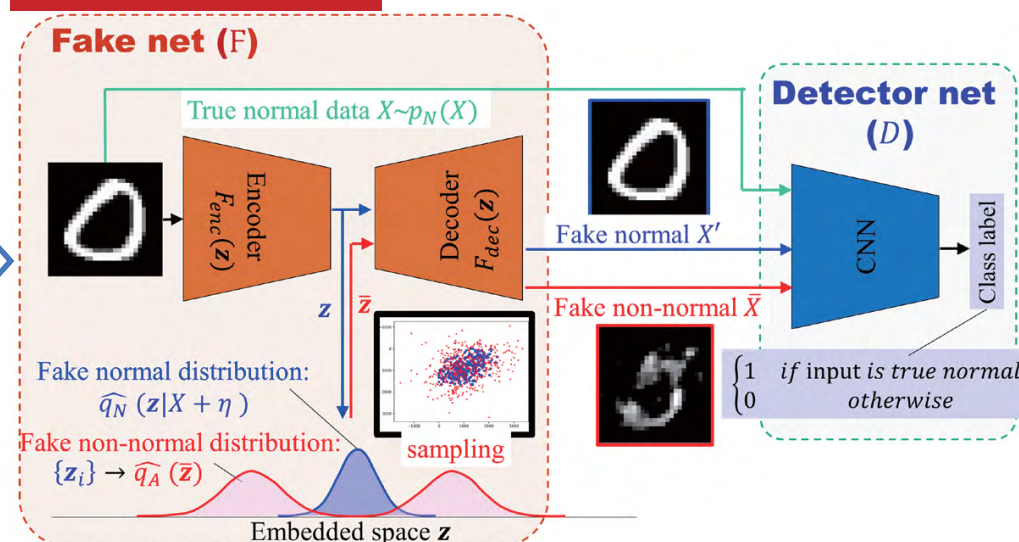
Problem

Trained D is for classifying X or X' but NOT for normal X or abnormal $\tilde{X}$.

Fake normal distribution $\hat{p}_N(X')$ True normal distribution $p_N(X)$



Our approach



Use the auto-encoder to generate pseudo abnormal data by sampling from the outside of the normal-data-region on the embedded space.

Bayesian inference on transcription elongation rates from total RNA sequencing

河村 優美 (総合研究大学院大学 統計科学専攻)

吉田 亮 (情報・システム研究機構 統計数理研究所)

目的

RNA ポリメラーゼ II (Pol II) の転写伸長プロセスを再構成

手法

状態空間モデルを用いたベイズ推論を行い, 逆問題を解く



<https://github.com/yoshida-lab/PolSter>

結果

RNAシーケンスの観測データからPol IIの存在確率を推定できることを実証

- Total RNA-seq: トランスクリプトーム計測の手法
- バイオインフォマティクスに基づく新規応用(方法論)の開拓
- Total RNA-seq とベイズ推論を組み合わせることで, ゲノム上の転写伸長速度を予測できることを示す
- これまで, 転写伸長速度の測定は薬剤処理した時系列データの移動距離と時間
- これは個々の実験デザインや膨大な時系列データの測定などコストがかかる
- 我々の手法は個別実験デザインは必要なくtotal RNA-seqの1回のデータだけ測定

ガウス埋め込みによるSentence Representation

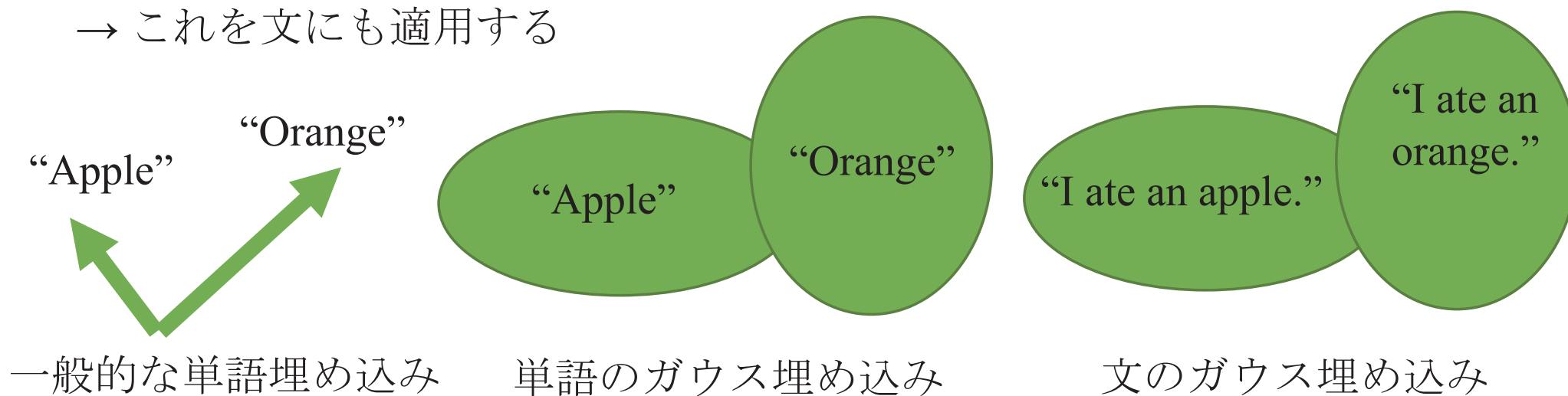
川田航希 (電通大) 西山悠 (電通大) 川野秀一 (電通大)

目的 文のより良い埋め込みを得るためにベクトルベースではない手法を提案

手法 単語の埋め込みで成功したガウス埋め込みを文の埋め込みに適用する

結果 STS Benchmarkにて従来法(Paragraph Vector)を上回る精度を達成

- 単語や文をベクトルで表現することで、類似度などを具体的に計算できる
- しかし、単語は様々な意味を持ち、文脈だけでは判断できない曖昧性もある
→ つまりベクトルだけでは不十分 → ガウス埋め込み
→ これを文にも適用する



Nonparanormal分布のグラフ構造に対するロバスト推定

奥田 諭史 笹井 健行 (DATUM STUDIO株式会社) 福島 真太郎 (株式会社トヨタIT開発センター)

目的

Nonparanormal分布の条件付独立性のグラフ構造のロバスト推定

手法

Nonparanormal分布に対してロバストなアルゴリズムである γ -Lasso を利用

結果

本研究の手法のロバスト性を数値的に検証した

正規分布



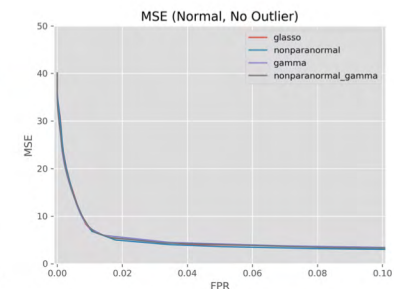
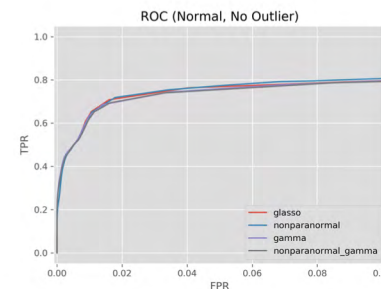
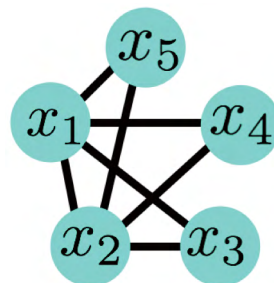
変換(輸送写像)

Nonparanormal分布
(正規分布のある種の一般化)

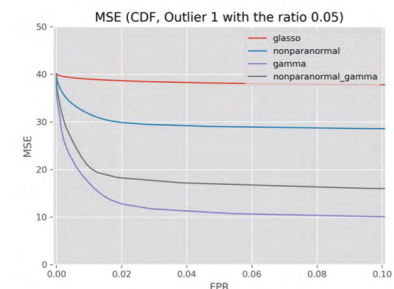
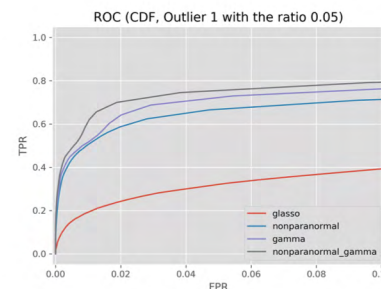


γ -Lasso

グラフ構造
(条件付独立性)



正規分布に対する実験



Nonparanormal 分布(CDF)に対する実験

階層的モジュール表現に基づく 多層ニューラルネットの推論構造の理解

渡邊 千紘 (NTTコミュニケーション科学基礎研究所)

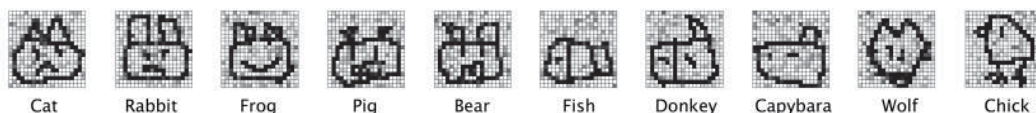
目的

多層ニューラルネットの学習結果を
解釈しやすい形で表現すること

手法

入出力次元との相関を元に各ユニットの特徴
ベクトルを定義し、階層的クラスタリングを適用
ニューラルネットの階層的クラスタ構造と
各クラスタの役割についての知識獲得を実現

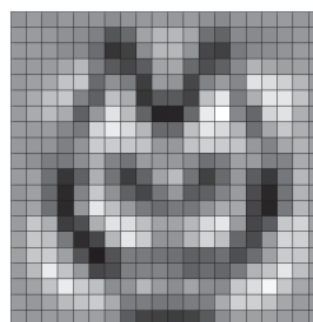
結果



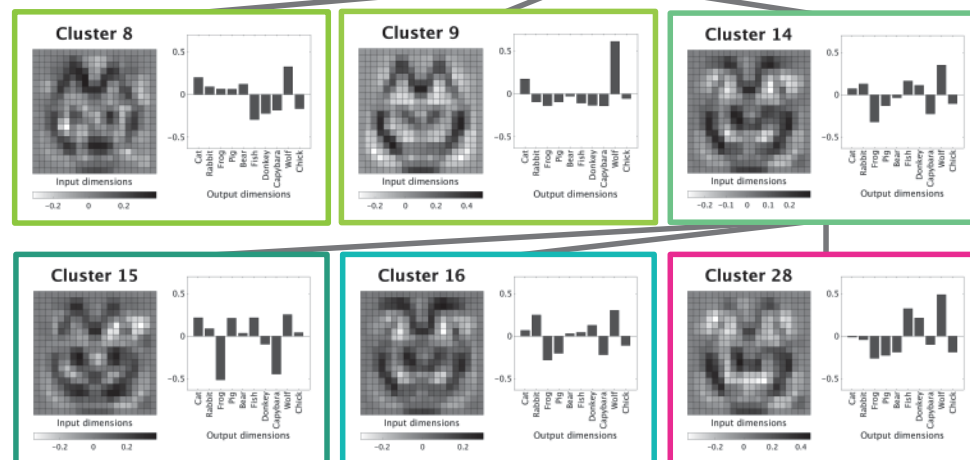
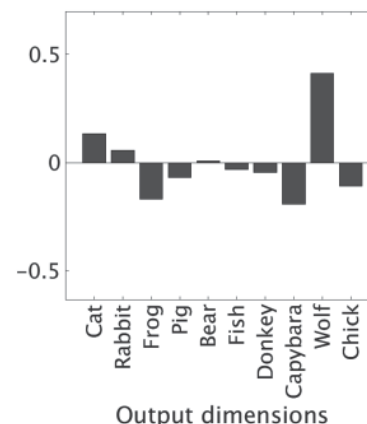
クラスタ
の役割

各入力次元との
相関の平均値
(どこを見て)

Cluster 5



各出力次元との
相関の平均値
(何を認識するか)



Covariate Selection in Propensity Scores for Binary Outcomes via Sparse Regularization

張 安杰 (電通大) 川野 秀一 (電通大)

目的

因果推論において傾向スコアを構成する共変量を選択する

手法

スパースモデリングに基づく共変量選択法を確立

結果

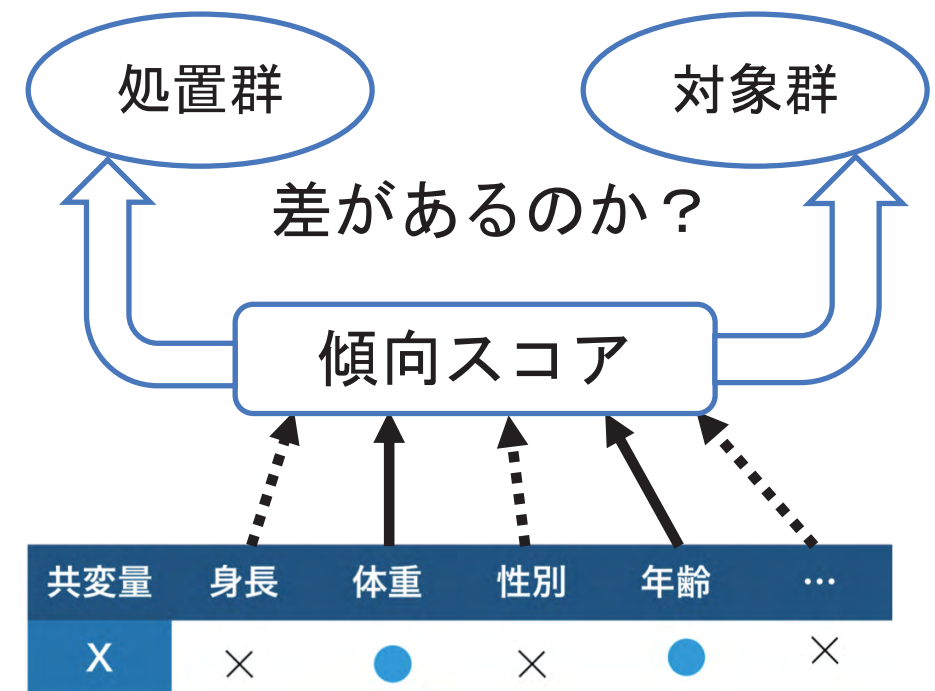
数値実験を通して因果効果の推定精度の向上・安定性が示された

背景

- 因果推論における因果効果推定において、傾向スコアがよく用いられる
- 傾向スコアは共変量により構成
- 近年共変量の個数は膨大

開発した方法論

従属変数が2値のときのスパースモデリングを用いた共変量選択手法



スパースモデリングによる共変量選択

VAEの事前分布の改善

高橋 大志 岩田 具治 山中 友貴 山田 真徳 八木 哲志 (NTT)

目的

VAEの潜在変数の事前分布を改善し、未学習を防ぐ

手法

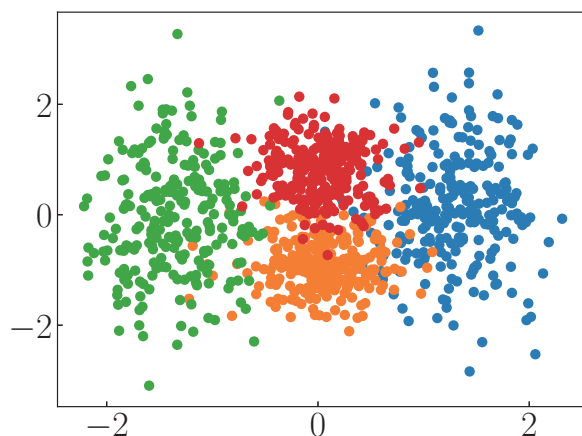
最適な事前分布を導入し、密度比推定を用いて計算可能にする

結果

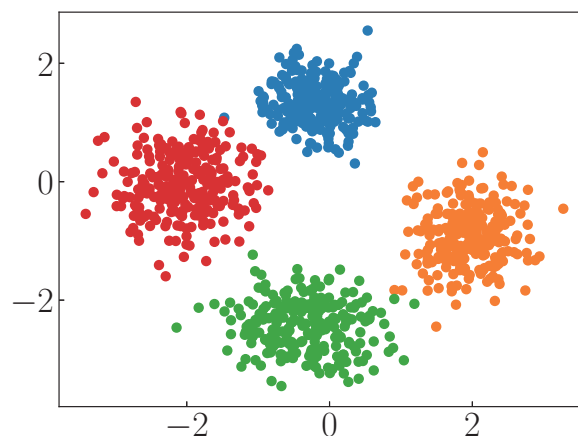
テスト尤度の向上を確認した

- 本発表では、VAEにとって最適な事前分布である**周辺事後分布**を導入する
- 従来の学習法は使えなくなるため、**密度比推定**を用いた学習法を提案する

例: 4次元One Hot Vectorに対する潜在変数



通常のVAE



提案手法

潜在変数がoverlapせず、
decodeしやすくなる



再構成誤差が減るため、
尤度が向上する！

Matrix Co-completion for Multi-Label Classification with Missing Features and Labels

Miao Xu¹, Gang Niu¹, Bo Han^{2,1}, Ivor W. Tsang², Zhi-Hua Zhou³, Masashi Sugiyama^{1,4}
¹RIKEN, ²UTS, ³Nanjing U., ⁴U.Tokyo

目的 Solve multi-label classification with missing features and labels

手法 Propose theoretical guaranteed matrix completion algorithm

結果 Achieve better Empirical Results on both synthetic and real data

$$\text{MC}\left(\begin{bmatrix} \mathbf{x} \end{bmatrix} \begin{bmatrix} \mathbf{z} \end{bmatrix}\right) \quad \begin{array}{l} \text{Assume} \\ \Pr(Y_{ij} = 1) = \sigma(Z_{ij}) \\ \sigma(\cdot): \text{sigmoid function} \end{array}$$

We propose an optimization objective involving both Frobenius norm and trace norm on matrix based on theoretical guarantee.

VAEとガウス過程による高次元データの圧縮と同時分節化

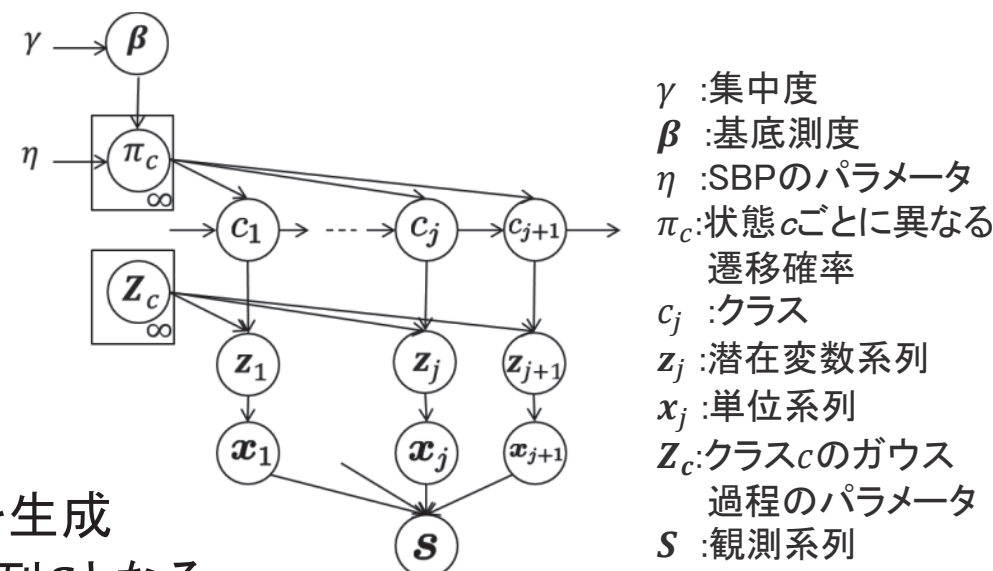
長野 匡隼 (電通大) 中村 友昭 (電通大) 長井隆行 (阪大) 持橋 大地 (統数研)
小林 一郎 (お茶大) 高野 渉 (阪大) 金子 正秀 (電通大)

目的 高次元の複雑な時系列データの教師なし分節化

手法 統計的に観測系列を単位系列に教師なしで分節化する手法 (HDP-GP-HSMM) にVariational Autoencoder(VAE)を導入

結果 教師なしで高次元時系列データを圧縮し、特徴を抽出しながら、単位系列へ分類・分節する

- 複雑かつ膨大なデータの分節化を統計的手法により教師なしで行う
- 以下のような生成プロセスを仮定
 - 階層ディリクレ過程より遷移確率 π_c を定めクラス c_j を生成
 - 単位系列 x_j を表現する潜在変数系列 z_j をガウス過程により生成
 - VAEのDecoderにより z_j から、単位系列 x_j を生成
 - 単位系列 x_j を結合することによって観測系列 S となる



Tree-LSTMの拡張による 抽象構文木のより良い分散表現獲得に向けて

*紫藤 佑介⁽¹⁾ 小林 靖明⁽¹⁾ 山本 章博⁽¹⁾ 宮本 篤志⁽²⁾ 松村 忠幸⁽²⁾ 嶺 竜治⁽²⁾

⁽¹⁾京都大学大学院情報学研究科 ⁽²⁾ (株) 日立製作所基礎研究センタ

目的

ソースコードの静的解析タスクにおいて木構造(=抽象構文木)を特徴抽出に利用することによる精度の向上.

手法

抽象構文木を扱う際に生じるTree-LSTMの問題を解消した拡張の提案.

結果

ソースコードのコメント自動生成タスクにおいてBLEU-4を用いた評価で既存手法を超える精度を獲得.

- Tree-LSTM[Tai+, 15]ではソースコードの特徴抽出に必要である, **不定数の子ノードの順序**を扱うことが困難である. 提案手法ではこれを解消した.



抽象構文木において順序は重要

評価指標	[Iyer+, 16]	[Hu+, 18]	Tree-LSTM	Ours
BLEU-4	.0852	.1830	.2034	.2099
GLEU-4	.1293	.2321	.2527	.2617

HMLasso: 高次元・高欠測データに対するスパースモデリング

高田 正彬 (東芝) 藤澤 洋徳 (統計数理研究所) 西川 武一郎 (東芝)

目的

高次元 & 高欠測のデータ. 回帰モデルを推定. 欠測補完しない.

手法

ペアワイズ共分散でデータを効率利用. **実測率で重み付け**して正定値化.

結果

推定誤差を大幅削減. 実測率重みを理論保証.

- 基本的な考え方: 回帰モデルの推定には共分散だけで十分. 欠測補完不要.

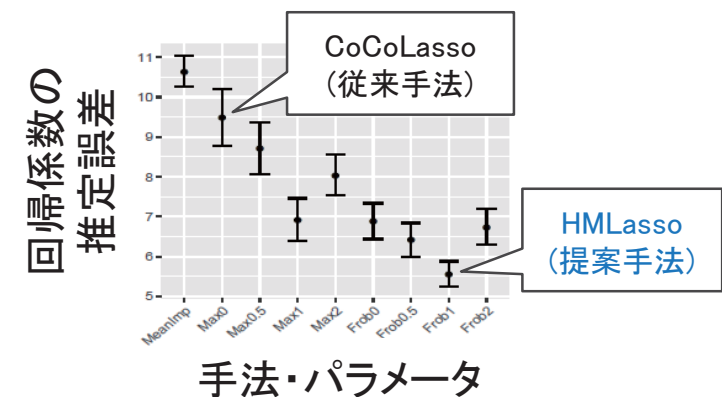
$$\hat{\beta} = \operatorname{argmin}_{\beta} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 = \operatorname{argmin}_{\beta} \beta^T X^T X \beta - y^T X \beta + \lambda \|\beta\|_1 \quad (X: \text{説明変数}, y: \text{目的変数})$$

- 従来手法: ペアワイズ共分散行列を正定値化 (CoCoLasso [Datta & Zou 17 AS])
- 提案手法: 共分散行列の正定値化にて**実測率で重み付け**.

$$\tilde{\Sigma} = \operatorname{argmin}_{\Sigma \geq 0} \|R \odot (\Sigma - S^{\text{pair}})\|_F^2 \quad (R: \text{実測率行列}, S^{\text{pair}}: \text{ペアワイズ共分散行列})$$

結果

- 数値実験にてCoCoLassoより**推定誤差43%減**.
- 実測率重みの最適性を証明** → ポスターにて



Designing Loss Function which Generate Latent Space for image in-betweening

Paulino CRISTOVAO(筑波大), Yusuke TANIMURA(産総研),
Hidemoto NAKADA(産総研) and Hideki ASOH(産総研)

目的

Design a Loss function which Generates image in-between using latent models.

手法

Based on Generative Models using Variational Autoencoders.

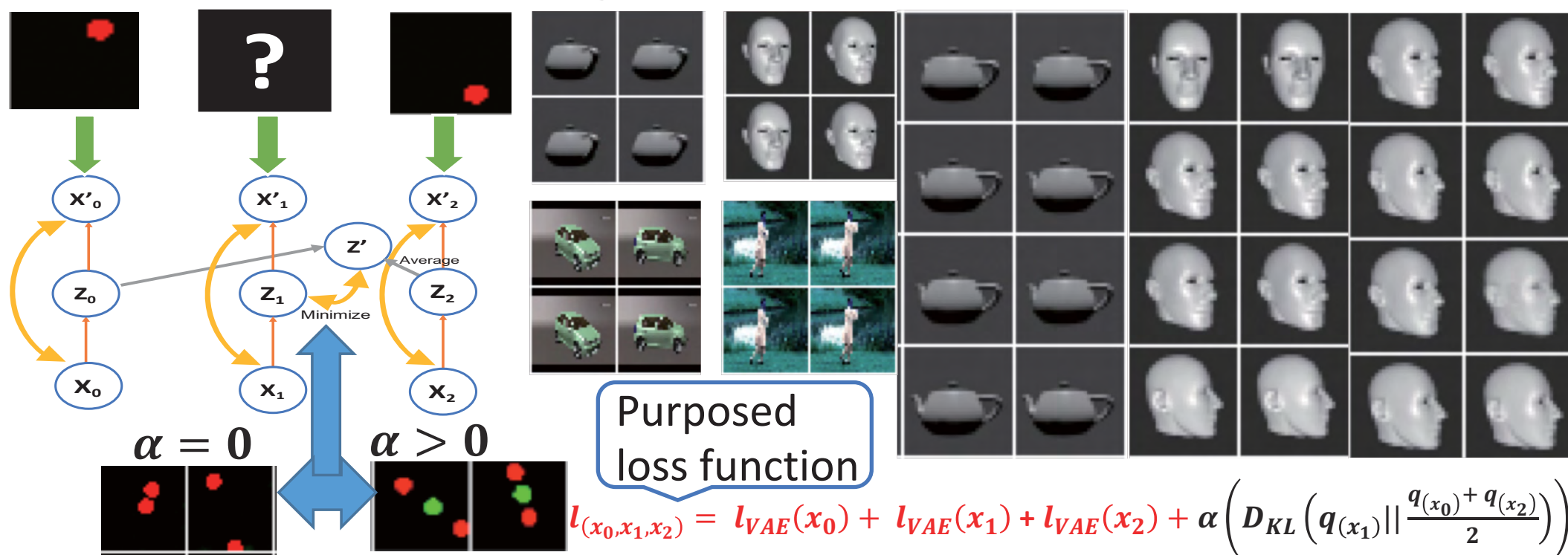
結果

The goal was achieved as the generated in-between image keeps the spatial location of the object as shown below.

Network Architecture

Image Reconstruction

In-between image



One-shot Learning using Triplet Networks with KNN

周 牧 (筑波大) 中田 秀基 (産総研) 谷村 勇輔 (産総研)

目的

To predict the query images by given only single example of each class.

手法

Train the augmented data on Triplet Network, then classify by KNN.

結果

Accuracy of the test dataset is 67.34% .

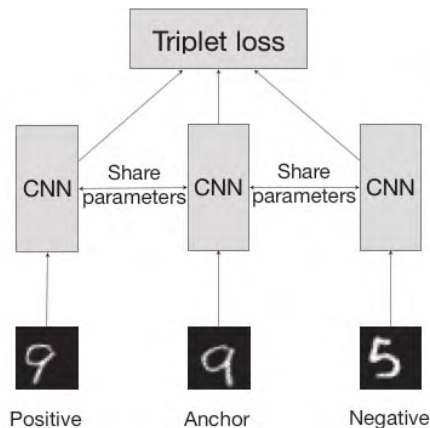


Figure 1. Triplet Network Model

Procedure

- Pick the data for training from MNIST dataset.
- Augment the data which has only 1 sample, then prepare the triplet pairs.
- Get the embeddings of training/test dataset after training on the triplet network.
- Classify the data with KNN classifier using the embedding vector.

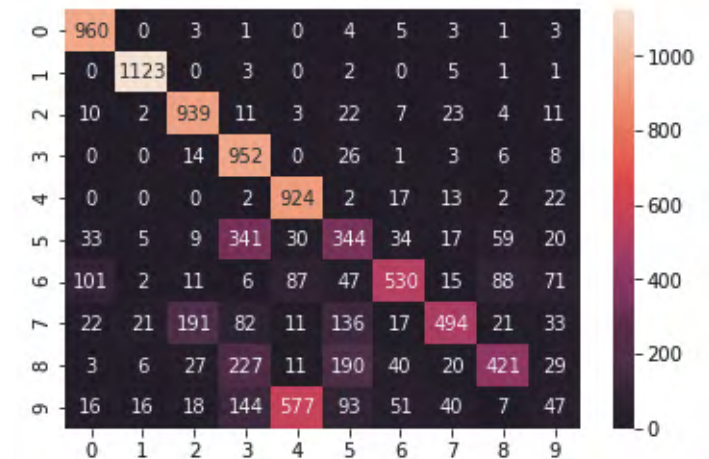


Figure 2. The confusion matrix of the result.

Kernel Gragh Lapracian Features

～ 重みつきデータを用いた局所性保存射影のカーネル化～

高橋 春輝(工学院大・情報) 竹川 高志(工学院大・情報)

目的

SN比が低く多様体構造をしたデータの前処理

手法

カーネル化したGragh Lapracian Features

結果

ノイズにより近づいたクラスタ間の分離を可能とした

- Gragh Lapracian Features(GLF)

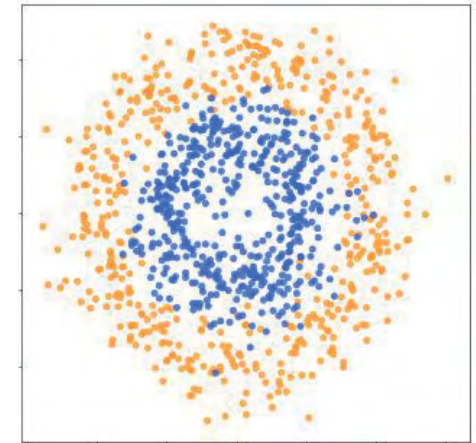
① 変換後のデータの分散を最大化

② 元のデータ間の類似度を基に変換後のデータ間の距離を決める

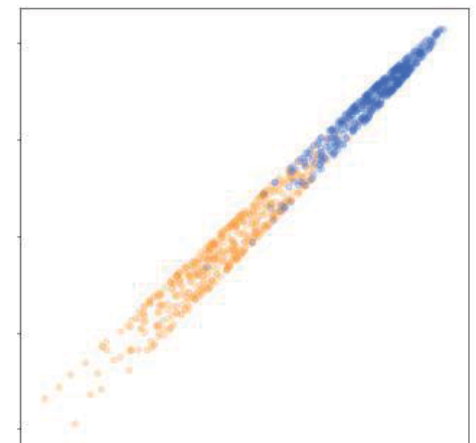
➤ ①②を条件に元データを線形変換

- Kernel Gragh Lapracian Features(KGLF)

➤ ①②を条件に特徴量データを線形変換



KGLF



観測量間の動的構造を抽出する ベクトル値再生核ヒルベルト空間における動的モード分解

藤井 慶輔（理研AIP） 河原 吉伸（阪大・理研AIP）

目的

観測量間で依存構造を持つデータに適用する動的モード分解の開発

手法

ベクトル値RKHSで定義されたKoopman作用素のスペクトル解析を定式化

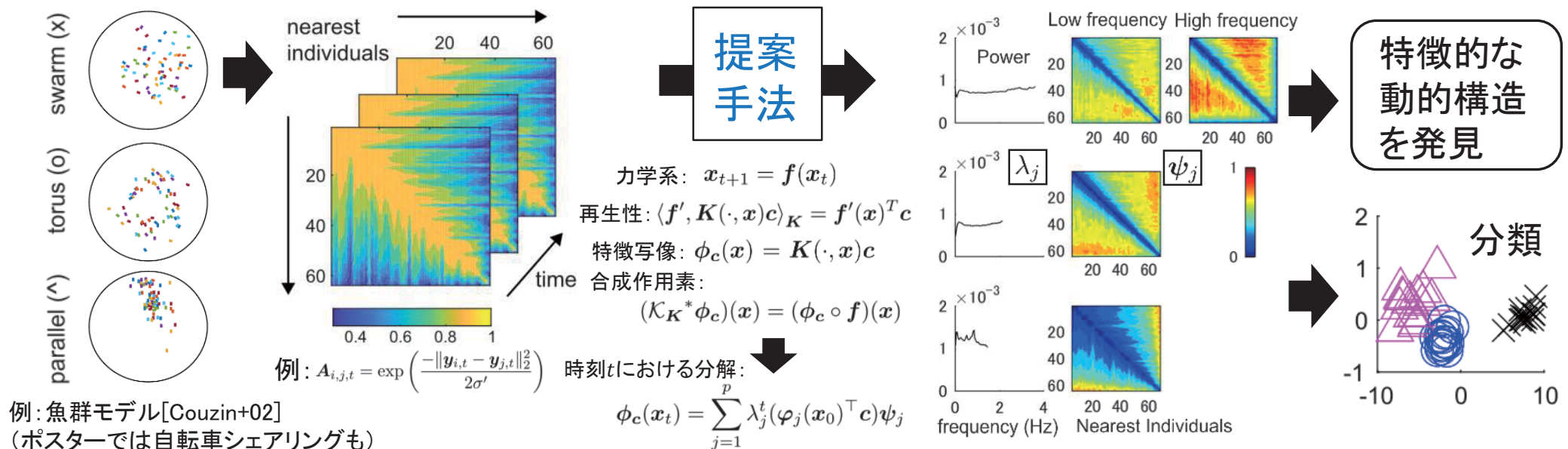
結果

上記のデータから特徴的な動的構造（時間特性や空間的コヒーレンス）を抽出

提案手法の全体像

（詳細はポスターまたは <https://arxiv.org/abs/1808.10551>）

1. データ収集 2. 入力系列の作成 3. 時間・空間モードに分解 4. 系の理解



(発表取り下げ)

目的

手法

結果

隠れ状態を持つ多腕バンディット問題における 方策の検討

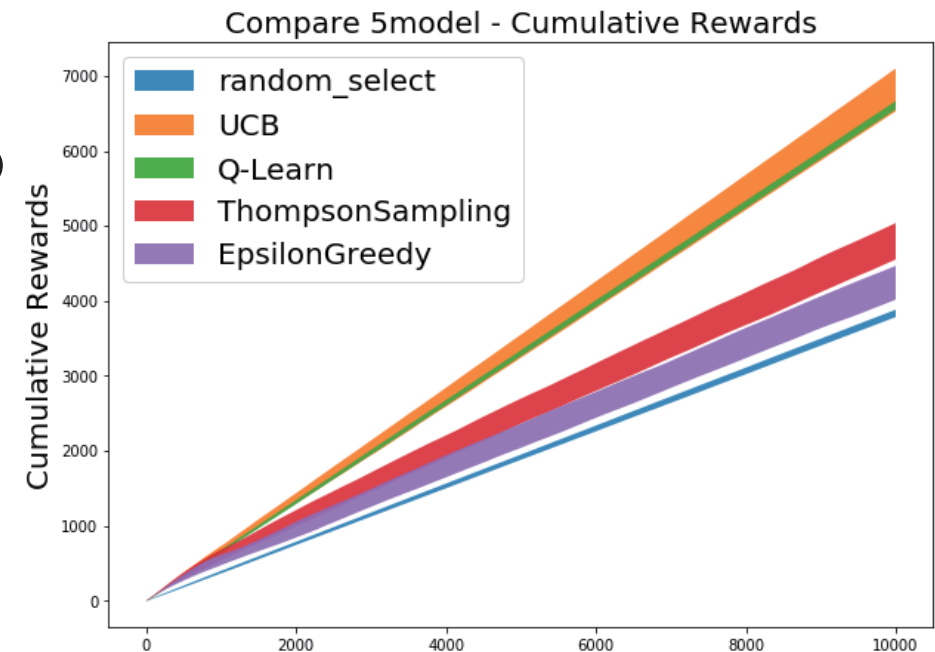
工藤 亘平（工学院大） 竹川高志（工学院大）

目的 人の行動に近いモデルでバンディット問題の報酬最大化

手法 確率的バンディット問題の方策に加えQ学習

結果 UCB方策, Q学習は他より多くの報酬を獲得

- 今回使用したモデル
一定期間ある状態に留まり、その後各アームの
当選確率が異なる別の状態に確率で遷移する
- 状態に留まる期間を短くした場合
獲得報酬は減少傾向にある
- 今後、状態推定をする方策を適用
例 Q-network, カルマンフィルタ



潜在的ディリクレ配分法を用いた 問合せ文書と回答文書の関係分析

大川 順也(早大) 雲居玄道(早大) 後藤正幸(早大)

目的

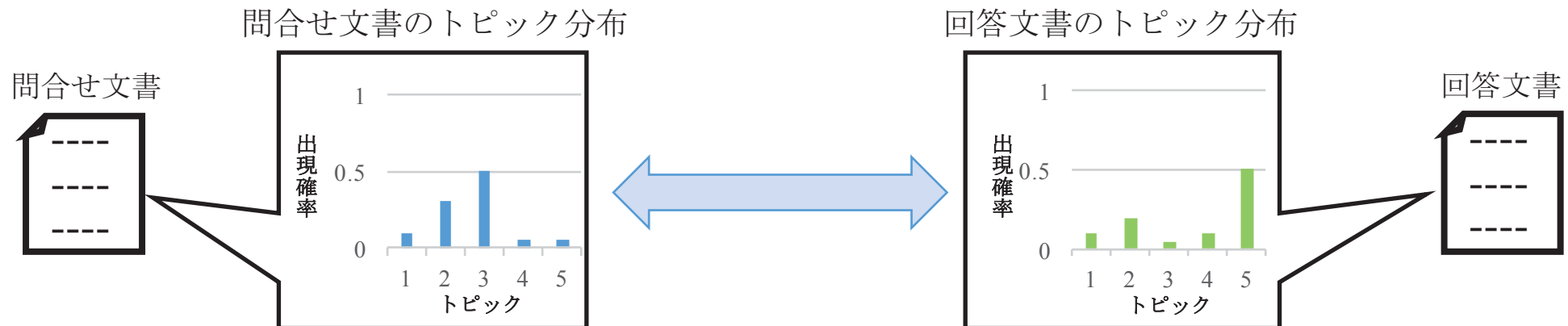
社内QAシステムの回答文書作成プロセスにおいて、問合せ文書の内容によって、過去の回答文書の提示で良いか、あるいは人による個別対応が必要かを分類

手法

問合せ・回答文書のそれぞれ潜在的ディリクレ配分法(LDA)を適用し、双方のトピック分布の類似度によって、問合せに対して回答を対応させる手法を提案

結果

実データへの適用により、問合せ文書のトピックが「請求処理」、「ログイン」、「インストール、権限」、「車両、端末」の場合について自動対応が可能となった



LDAにより推定したトピック分布を比較することで、双方のトピックの関連度を算出

問合せ文書のトピックによって、
自動対応が可能か、人による個別対応が必要かを分析・考察

共変量シフトの重み関数選択による因果効果推定

横山 悠 (京都大学) 下平 英寿 (京都大学/理研AIP)

目的 重み付き回帰による因果効果推定法のための重み関数を選択する

手法 共変量シフトの視点を導入 → 因果効果推定量の漸近分散に注目

結果 従来法と比較して分散の小さい因果効果推定量が得られる

Step 1. 重み関数を決定

【従来法】

重み関数は固定

☹ 説明変数の分布によっては
最適でないことも

【提案法】

最適な重み関数を選択

- 因果効果推定量の漸近分散を最小化
- 「テストデータの最適化」と解釈できる



Step 2. 重み付き回帰を実行



Goal: 平均因果効果

コヒーレンスが大きい線形観測による LASSOの近似的性能評価

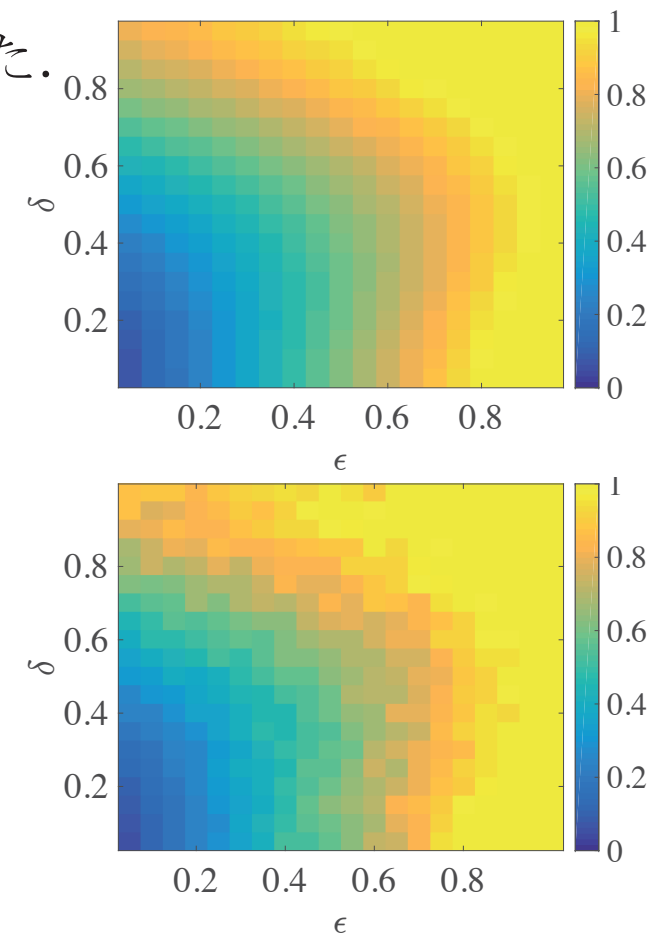
井原 みのり 岩田 一貴 末松 伸朗 三村 和史
(広島市立大学)

目的 観測行列のコヒーレンスのスパース推定への影響を調べる.

手法 統計物理的な手法でLasso推定量の平均二乗誤差を評価する.

結果 推定性能のコヒーレンス依存性が明らかになった.

- Lasso推定量 $\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} (\|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1)$
- $\mathbf{y} = A\mathbf{x}_0 + \mathbf{n}$: 観測
- $A = (a_{ij}) \in \mathbb{R}^{M \times N}$, $a_{ij} \sim \mathcal{N}\left(\frac{1}{\sqrt{M}} \times \frac{h}{\sqrt{1+h^2}}, \frac{1}{M} \times \frac{1}{1+h^2}\right)$
: 既知の観測信号
- $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 I_N)$: ノイズ, δ : 圧縮率, ϵ : 信号密度
- Lasso推定値の誤差 $\text{MSE} = \mathbb{E}_{A, \mathbf{x}_0} (\|\hat{\mathbf{x}} - \mathbf{x}_0\|_2^2)$
- 理論 (上) と実験 (下) の比較
- $\lambda = 0.1, h = 1.0, \sigma^2 = 0.3$ の場合



形状特徴項を用いた VAEの形状復元性能の向上

佐藤 玄† 前田 宗彦† 鈴木 一将† 内田 雄太† 大島 誉寿† 米倉 一男†
† 株式会社IHI

目的 外観検査におけるVAEの復元性能の向上

手法 VAEの復元誤差に、交差エントロピーによる誤差に加え、**形状特徴項を追加**

結果 欠陥などの未学習形状を入力した場合でも特徴を維持した復元が可能

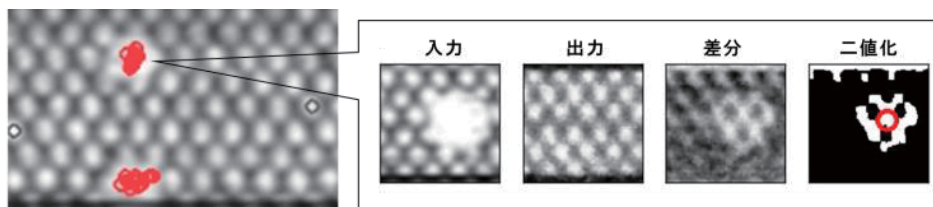


図1 VAEを用いた欠陥検知の概念図

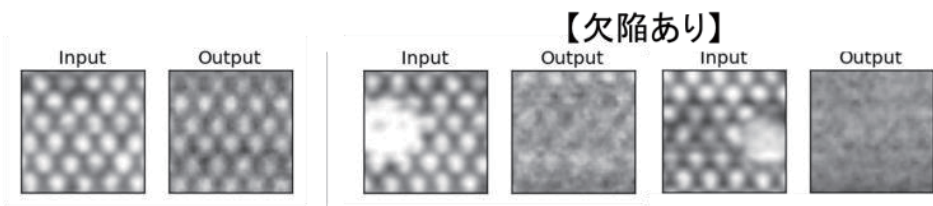


図2 欠陥を入力した際に学習した特徴が失われた例

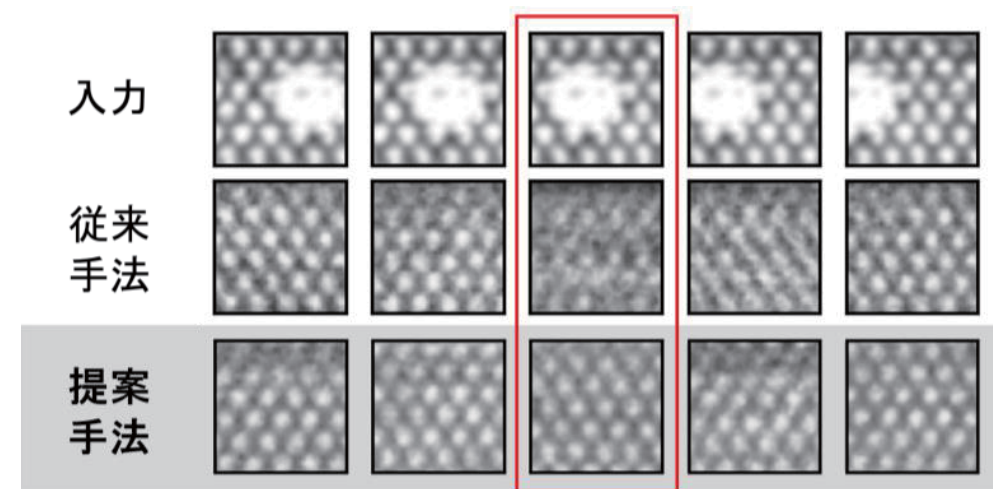


図3 提案手法と従来手法との比較

強推移性を利用した漸近一致性を持つ 大規模ベイジアンネットワーク構造学習

本田 和雅 (電通大) 名取 和樹 (電通大) 磯崎 隆司 (電通大) 植野 真臣 (電通大)

目的

これまでできなかった大規模のベイジアンネットワーク学習について
漸近一致性を持つ高速な構造学習アルゴリズムの提案

手法

独立性の性質である強推移性を利用することで独立性検定の回数を削減

結果

提案アルゴリズムによる小規模ネットワークにおける真の構造の推定と
大規模ネットワークにおける従来手法より高速な学習の実現

- 強推移性における左辺の独立性から右辺の独立性を発見する定理を証明

[定理]

$G = (\mathbf{V}, \mathbf{E})$ を非巡回有向グラフとし, $X, Y \in \mathbf{V}$ とする. また, ノード Y は
ノード X の非子孫とする. このとき, $\forall A \in \mathbf{V} \setminus (\{X\} \cup \{Y\} \cup \mathbf{Pa}(X, G) \cup \mathbf{W})$ で

$$I(X, Y \mid \mathbf{Pa}(X, G)) \Rightarrow I(X, A \mid \mathbf{Pa}(X, G)) \text{ or } I(A, Y \mid \mathbf{Pa}(X, G)).$$

$\mathbf{W}$: X の子孫で, X と Y を結ぶ路で合流結合するノードとその子孫からなる集合

$\mathbf{Pa}(X, G)$: X の G における親ノード集合

On Training a Joint Magnitude and Phase Deep Generative Model

Aditya Arie Nugraha (理研AIP) 関口 航平 (理研AIP・京大)
吉井 和佳 (理研AIP・京大)

目的

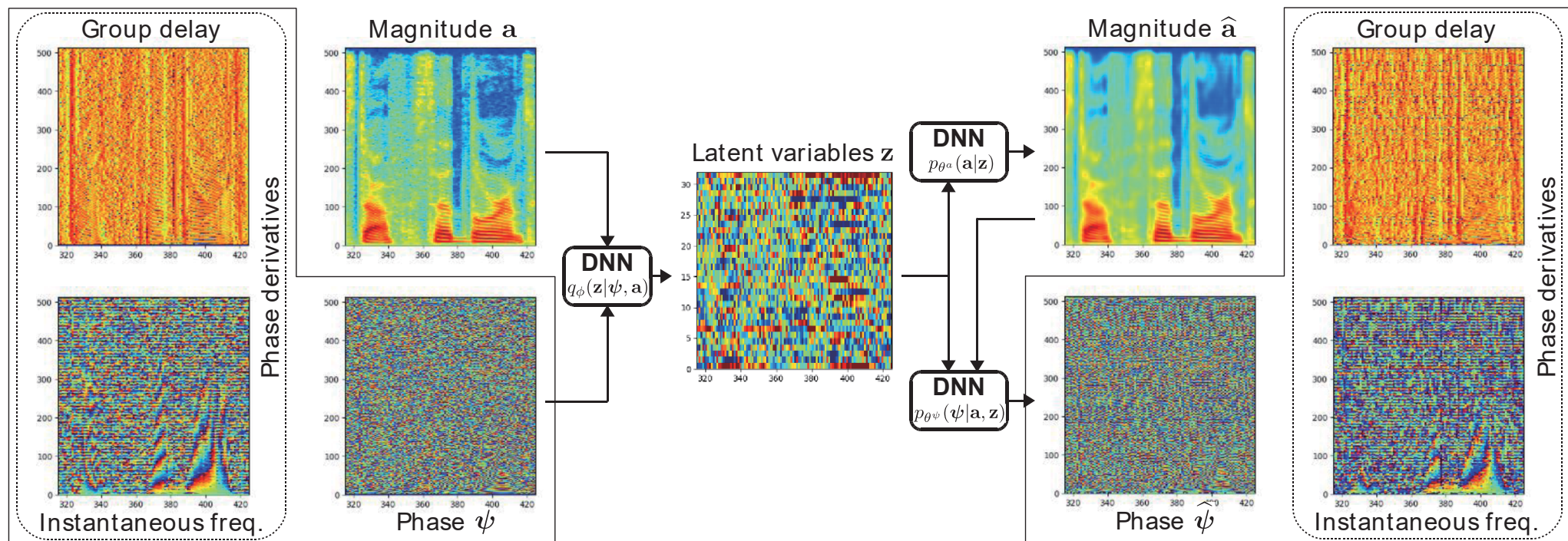
To obtain a generative model of the speech complex spectrograms

手法

Training a VAE to model magnitude and phase, with phase derivatives as constraints

結果

Achieved a good speech reconstruction with a high intelligibility



Group delay: the derivative along the frequency axis -- Instantaneous frequency: the derivative along the time axis

Learning Algorithms in Molecular Adaptation: Adaptive response of side-chain dynamics in binding site of ribonuclease inhibitor

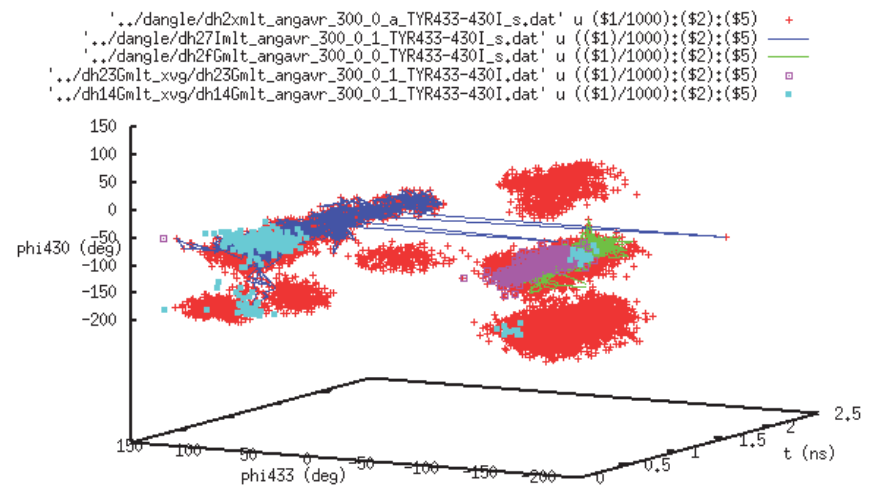
Nobuyuki Takahashi (Hokkaido University of Education, Hakodate)

目的 Temporal tuning of the protein binding for synaptic maintenance.

手法 Distance and orientation analysis of time dependent molecular structures.

結果 Conformational relaxations of side chains change binding response.

- RNA metabolism for long term memory formation needs “temporal tuning”.
- Atomic interactions take a role in feature extraction and the recognition based on the H-bonds of side chains.



Internal bond rotational angles of the TYR433 and TYR430 side chains were plotted as a function of time.

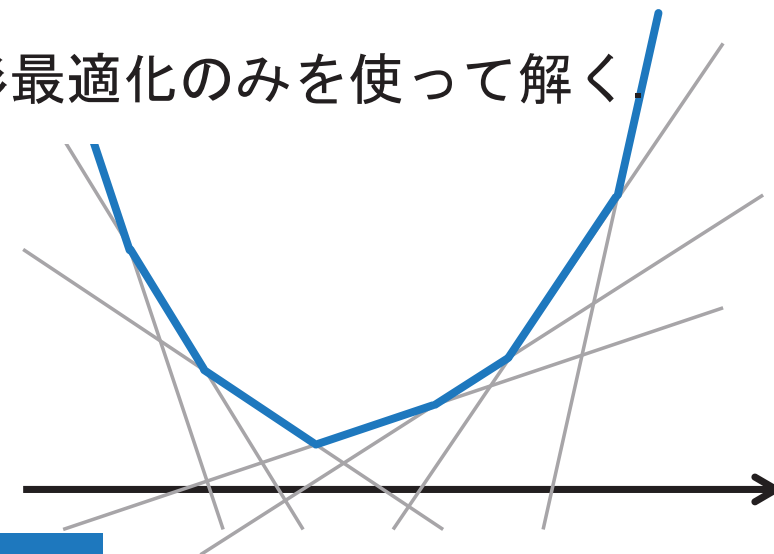
Frank-Wolfe型ブースティングの高速化

三星 諒太郎(九大) 畑埜 晃平(九大・理研AIP) 瀧本 英二(九大)

目的 1-ノルムソフトマージン最適化問題を線形最適化のみを使って解く

手法 オンラインFrank-Wolfe法[Garber+, '13]に基づくブースティング手法.

結果 以下参照.



ε : 最適解との誤差
 N : データの次元数

	ERLPBoost	本研究	LPBoost
反復回数 の上界	$O\left(\frac{1}{\varepsilon^2} \log N\right)$	$O\left(\frac{N^2}{\varepsilon^2}\right)$	未知
反復毎の 部分問題	凸最適化 (遅い)	線形最適化 (高速)	線形最適化 (速い)

Remark

- 計算の速さと, 理論保証を両立.
- オンライン凸最適化手法の新たな応用例.

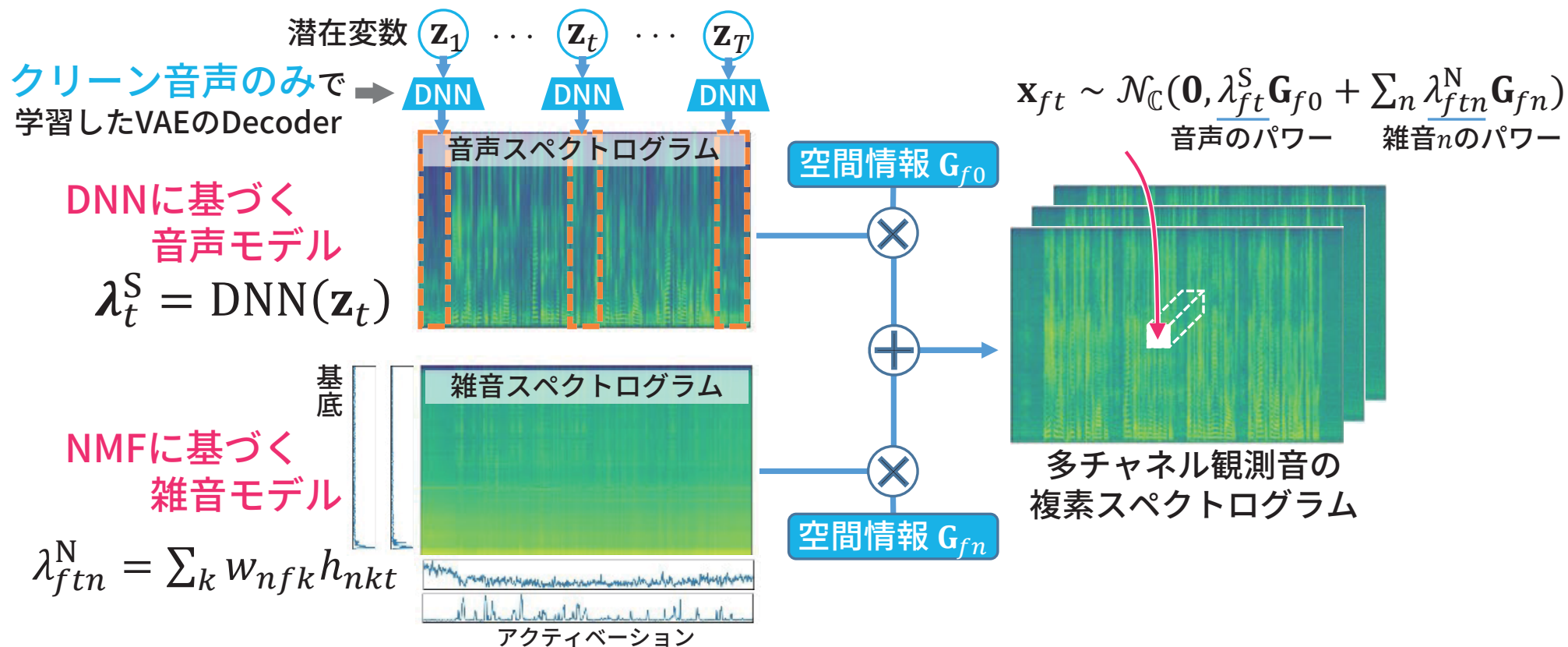
深層生成モデルを用いた半教師あり多チャネル音声強調

関口 航平 (理研AIP・京大) 坂東 宜昭 (AIST) Aditya Arie Nugraha (理研AIP)
吉井 和佳 (理研AIP・京大) 河原 達也 (京大)

目的 雑音データを学習に用いない環境依存性の低い音声認識

手法 VAEを用いてClean音声のみから音声スペクトルの生成モデルを学習

結果 雑音を動的に推定し音声のみを学習に用いる音声強調手法の開発に成功



単語分散表現に対する球面クラスタリングの適用

三橋 響(早大) 井上 真郷(早大)

目的 線形分離しやすい文ベクトルの作成

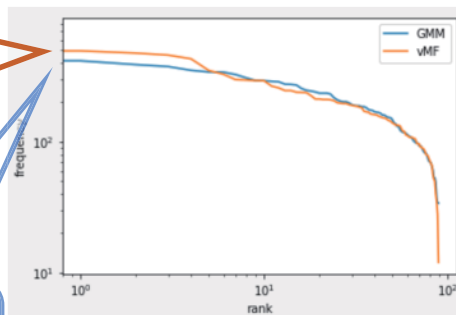
手法 movMF(球面クラスタリング)の結果を用い単語ベクトルを拡張する

結果 GMMを用いた際と分類精度にほぼ変化なし(微改善)

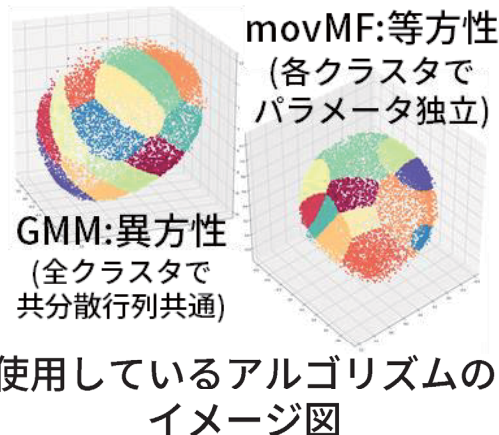
- word2vec[Mikolovら,2013]の単語ベクトルは球面クラスタリングが相性良い.
- SCDV[Mekalaら,EMNLP2017]では,共分散行列共通のGMM(パラメータ数約32000)でクラスタリング.
これをmovMF[Banerjeeら,2003](パラメータ数約12000)に変更.
- クラスタリング結果には差が見られた. RandomForestでの低次元部分空間で改善が見られた. 議論求む.

movMF:
最大724単語
windows,dos,ms,
disk,machine...

GMM:
最大492単語
david,john,mark,
bill,michael...



各クラスに属する
単語数分布



拡張に用いる 手法	線形SVM	Random Forest (100ノード)	Random Forest (1000ノード)
GMM	16.33	27.74	22.23
movMF	16.08	22.32	21.52

20Newsgroupsの誤答率[%]

単語の極性を埋め込んだ分散表現

中村 拓（東北大） 田 然（AIST） 乾 健太郎（東北大・理研AIP）

目的 単語の文脈的類似性・対立する極性の両方を表せる分散表現の実現

手法 単語の周辺文脈の予測と同時に、極性を**複数の**識別面で分離する

結果 極性を分ける複数の軸があり、それらを学習できることを確認した

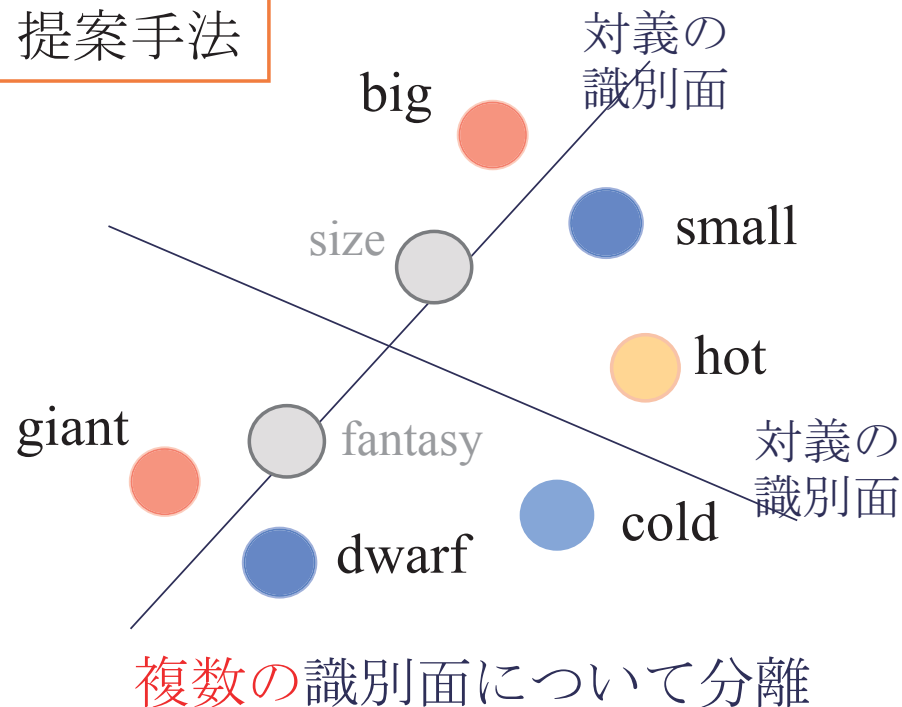
背景

- ・ 出現文脈の似た単語は近い分散表現になる
- ・ 対立する極性（**対義**や感情極性）の区別が難しい

課題

- ・ 対義語を引き離すと文脈的類似性に影響
- ・ 対義を分ける軸は**1つとは限らない**

提案手法



雑談チャットボットのための記憶注意付きSeq2Seqモデル

伊藤 駿汰 (早大) 井上 真郷 (早大)

目的

応答の根拠となる記憶情報を入力に加えることでより良い応答を生成

手法

より重要な記憶情報を強く反映する記憶注意機構とSeq2Seqの組み合わせ

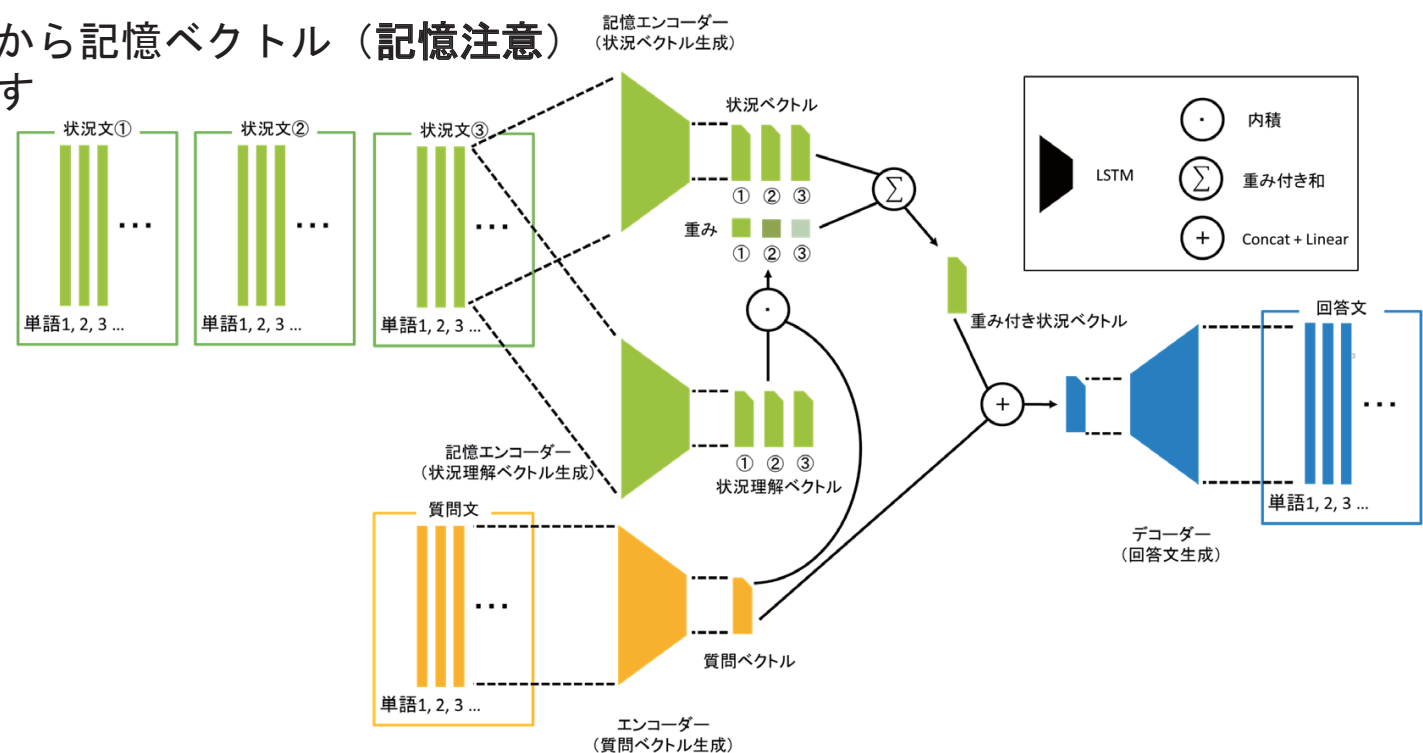
結果

学習が極めて不安定だが、部分的に記憶情報を反映した応答を生成

- 「文脈的に適切な応答」は過去の発話やプロフィール情報（記憶情報）を応答に反映する必要がある
- 応答の根拠となる記憶情報から記憶ベクトル（**記憶注意**）を生成し、デコーダーに渡す

課題

- 評価が困難
- 学習が極めて不安定、議論求む



Word2Vec における負例 Sampling の情報理論的一考察

竹中 誠 (首都大学東京) 小町 守 (首都大学東京)

目的

負例 Sampling におけるノイズ分布の意味をコーパスの特性に見出したい

手法

KLダイバージェンスに基づき近似されたノイズ分布を評価する

結果

良いノイズ分布は単語の多様性と共起単語の乱雑さのバランスで決まる

- Word2Vecにおける負例Samplingのノイズ分布
 - 頻度スムージングのためユニグラムの相対頻度の α 乗近似が用いられる
- α 乗近似による単語-文脈語の相互情報量欠損 $\Delta MI(X,Y)$

$$p(x, y) = p(x)p(y|x) \approx p(x)p(y)_\alpha \quad p(y)_\alpha = \frac{\#y^\alpha}{\sum \#y^\alpha}$$

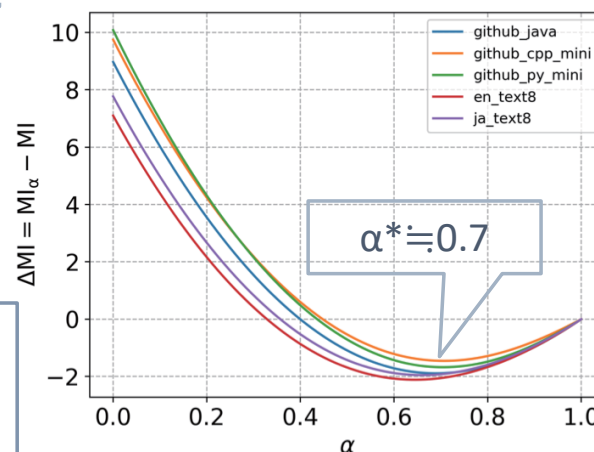
真のノイズ分布
(バイグラム分布)

近似されたノイズ分布

$$\begin{aligned} KL[p(x, y) || p(x)p(y)_\alpha] &= \sum_{x, y \in V} p(x, y) \log \frac{p(x, y)}{p(x)p(y)_\alpha} \\ &= MI(X, Y) + (\alpha - 1)(H_1(Y) - \sum_Y P(Y)H_\alpha(Y)) \end{aligned}$$

KL 最小の α を良い $= \alpha^*$ とすると、 α^* は単語分布の曖昧さ (H_1) と単語の多様性 (H_α) のバランスで決まる
(H_1 は Shannon entropy, H_α は Rényi entropy)

自然言語と人工言語の ΔMI



α^* はコーパスにおいて平均的に共起する単語がどの程度決まっているかを表す量と考えられる。すなわち、 α^* が 0 に近いコーパスは、共起単語がバラバラであり、 α^* が 1 に近いコーパスは共起単語に偏りがあるコーパスを意味する。

微分不可能な2段階最適化を用いたハイパーパラメータ学習について

奥野 貴之 (理研AIP) 武田 朗子 (東京大・理研AIP)
川名 哲裕 (東京工業大)

目的 ℓ_q 関数を含む正則化モデルのハイパーパラメータを効率的に決定したい

手法 2段階最適化法と平滑化法を組み合わせた反復法を適用

結果 点列がSB-KKT条件という最適性条件を満たす点に収束することを証明
ベイズアン最適化よりも高速に, かつ精度が高いパラメータを選択できた

ℓ_q 関数を含む正則化モデルの例

$$0 < q \leq 1$$

$$(1) \min_w g(w) + \lambda \sum_{i=1}^n |w_i|^q$$

ハイパーパラメータ

- $g: \mathbb{R}^n \rightarrow \mathbb{R}$: 平滑

Q) λ によってモデルの解 w_λ は異なる.
 λ をどうやって選ぶ?

Ans) テスト関数 $f(w_\lambda)$ の良し悪しで決定

2段階最適化問題

$$\min_{\lambda, w_\lambda} f(w_\lambda) \text{ s.t. } w_\lambda \text{ solves (1), } \lambda \geq 0$$

でも,これはどうやって解く?

測定困難なデータに対する 解析モデルベースの予測法の提案

草地 嵩（工学院大先進工） 高羽 洋充（工学院大先進工）

目的

電池のようなデバイスの測定困難なデータ予測のための、
解析モデルベースによる予測法の提案

手法

数値解析モデルによって教師データを生成
これを学習させ、測定困難なデータを予測する。

結果

Li-Air電池の対象とし、放電容量曲線から微細構造予測の精度を検証した。
実用に向けていくつかの課題項目がある。

電池の最適化に必要な 測定困難なパラメータの推定

- 空隙率
- 屈曲度
- 溶解度
- 拡散係数
- 反応速度定数
- etc...

電子顕微鏡、
化学分析など
破壊的な分析
が必要となる。

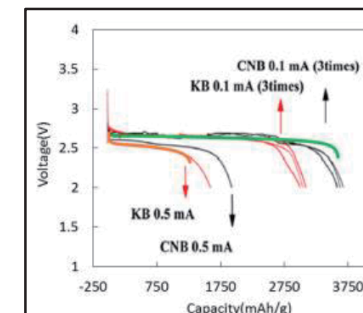
数値解析モデル



測定困難な
パラメータ

数値解析

測定容易な特性



機械学習

擬尤度の辺に対する最大化を通じた グラフィカルモデルの構造推定

高品 佑也(早大) 中谷 秀洋(サイボウズ・ラボ) 井上 真郷(早大)

目的

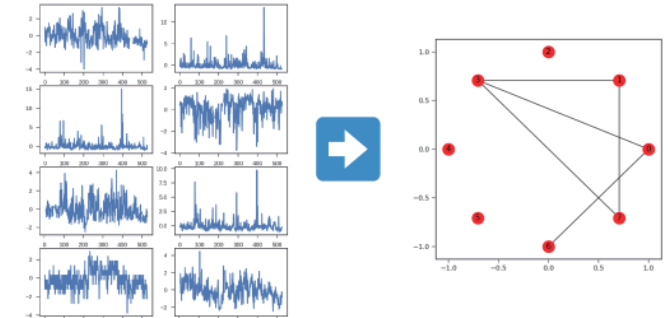
相互情報量を用いた Markov 確率場の構造推定

手法

擬尤度 (pseudolikelihood) の辺に関する最大化

結果

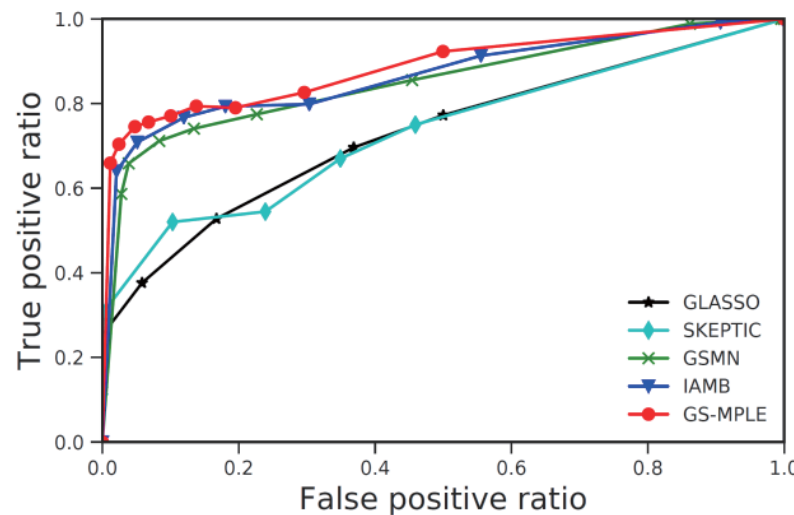
複数人工データで既存手法の精度を上回った



- 確率変数 X が $G = (V, E)$ に対してマルコフ確率場を成すとき:

$$\min_E \sum_{i=1}^D I(X_i; X_{V \setminus N[i]} \mid X_{N(i)}), \quad \text{where } \begin{cases} N(i): i \text{ の開近傍,} \\ N[i]: i \text{ の閉近傍.} \end{cases}$$

- 実験結果:



連続 HMM 構造推定の ROC 曲線

● : 提案手法

Online Intention Learning

Yasuhisa Suzuki(NEC), Wemer Wee(NEC)

目的

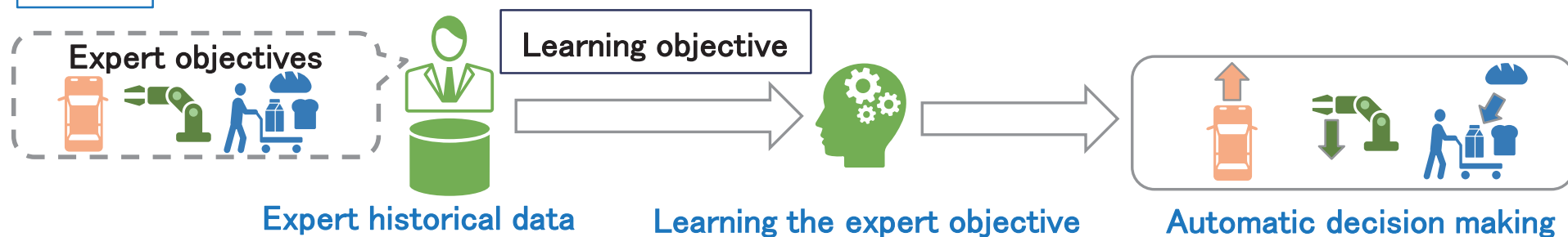
Infer an expert's objective function from observed decisions in online manner

手法

Probabilistic approach to online Inverse Optimization

結果

Accurate and efficient learning of the expert's objective parameters



- We are given observations of parameters and expert decisions over time: $\{(p_t, x_t)\}_{t=1}^T$.
- Assuming the decision x_t is a solution of an optimization problem of the form

$$\max \theta_e^T f(x) \text{ s. t. } x \in X(p_t),$$

we want to estimate the expert's parameter θ_e .

- Method: Probabilistic approach to online Inverse Optimization
 - Benefits: More “accurate” and faster learning & more complex objectives are handled.

partial AUC最適化を用いたマルウェア感染端末の分類

西山 泰史 熊谷 充敏 神谷 和憲 谷川 真樹 (NTT)

目的

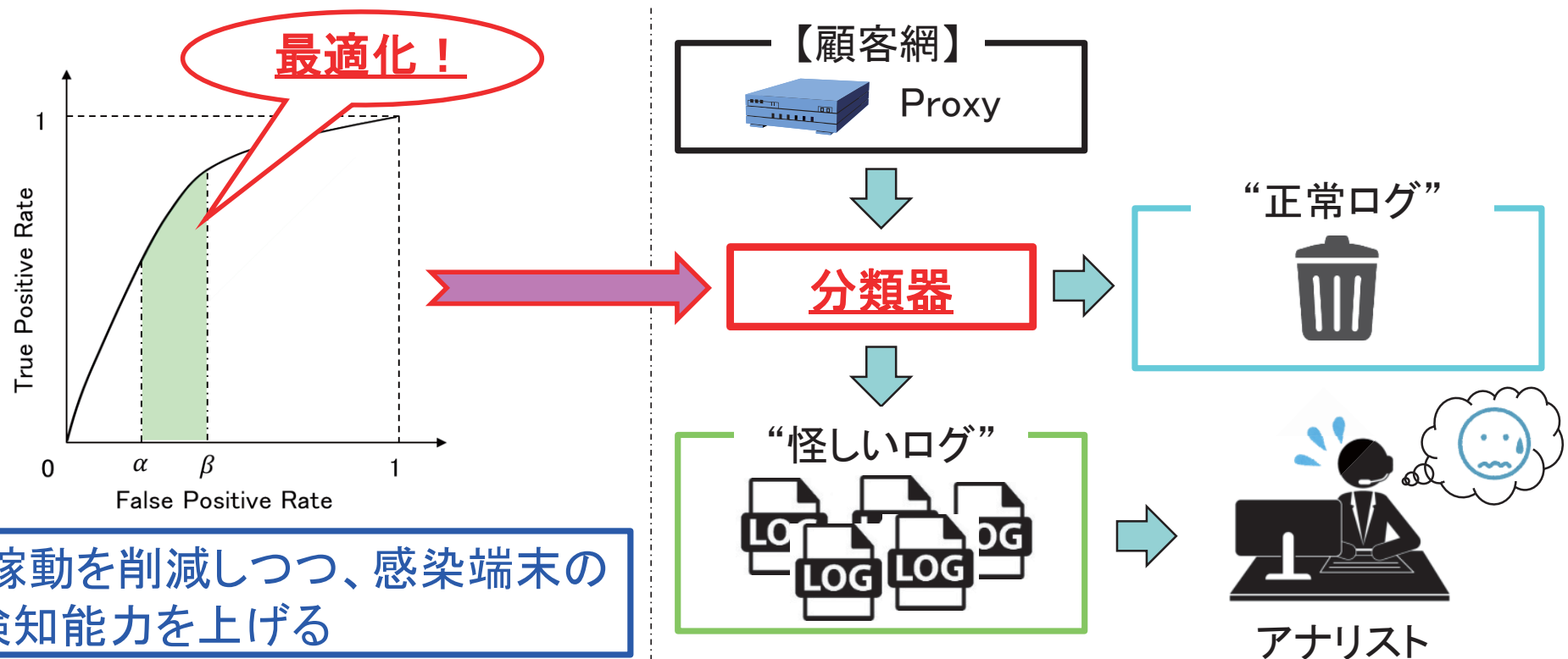
誤検知率を極力抑え、マルウェア感染端末の検知率をあげる

手法

寄与率算出/計算時間に配慮したpAUC最適化学習

結果

実在のHTTPログに提案法/先行研究/教師あり学習を実装し、効果を確認



パレートフロントに対するベジエ単体を用いた小標本近似法

小林 健 (富士通研・理研AIP) 濱田直希 (富士通研・理研AIP) 三内顕義 (理研AIP・慶大)
田中章詞 (理研AIP・慶大) 坂内健一 (慶大・理研AIP) 杉山将 (理研AIP・東大)

目的

単体的な多目的最適化問題の高次元パレートフロントの近似

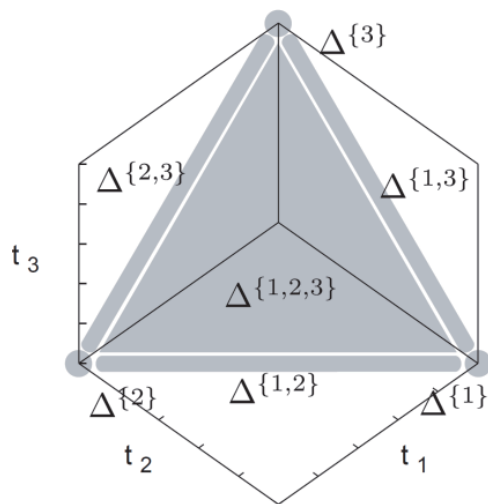
手法

ベジエ単体モデルによるパレートフロントの構造を活用した近似法を提案

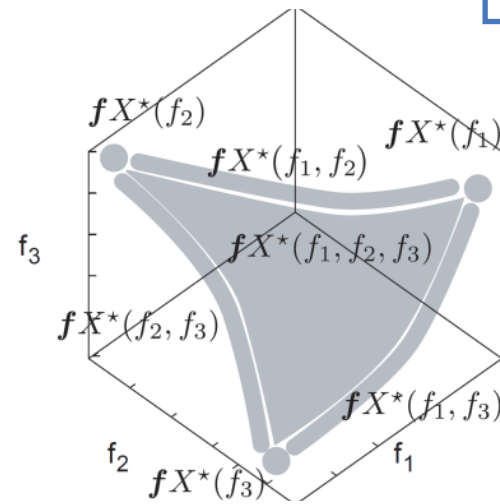
結果

ベジエ単体モデルの近似定理を証明/小標本による高精度近似を実験で確認

- 単体的な問題 (実問題で頻出):
パレートフロントが単体と同相



2次元単体



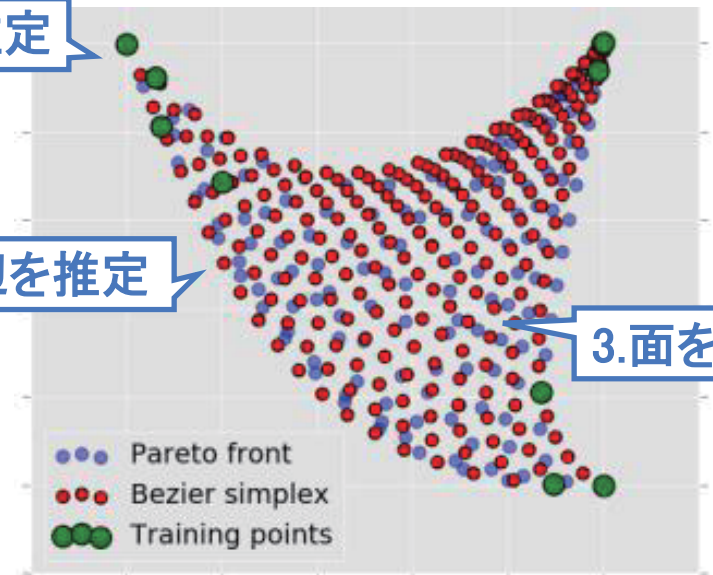
3目的最適化問題の
パレートフロント

- 提案法: **ベジエ単体**による近似

1.頂点を推定

2.辺を推定

3.面を推定



低次元単体から推定することで
パレートフロントの境界を考慮

温度付きSoftmax関数を含む 深層学習器の実験的性能評価

岡田 莉奈 長谷川 聡 (NTTセキュアプラットフォーム研究所)

目的

手法

結果

機械学習に基づく ペプチド由来イオンピークの新規検出手法

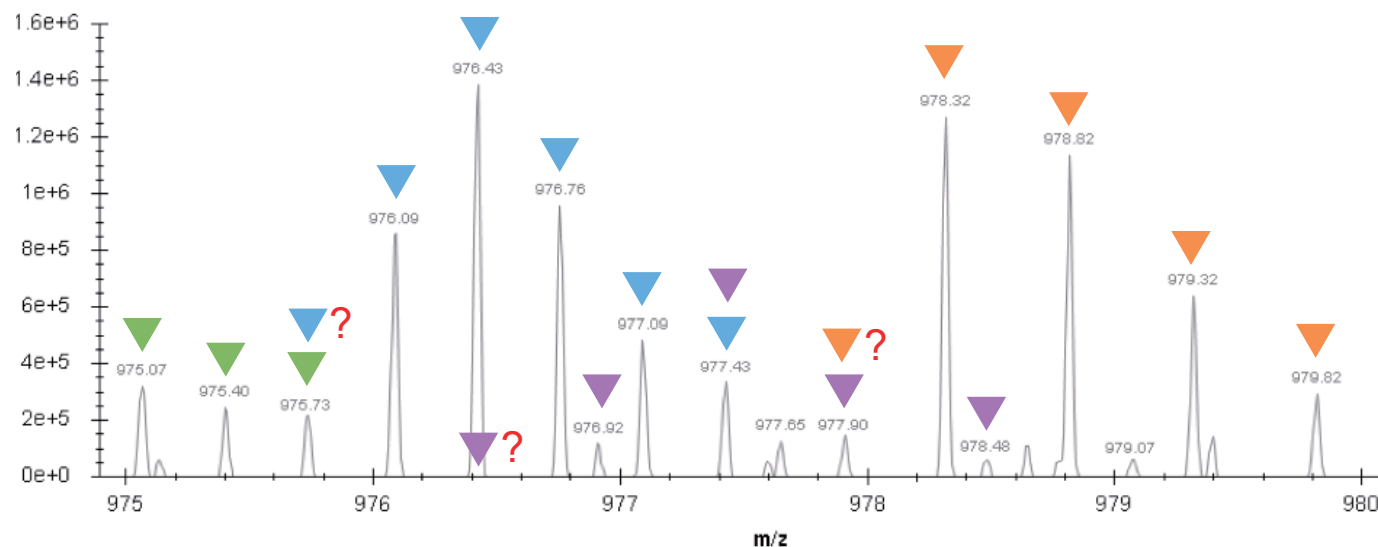
守屋勇樹 (DBCLS) 田畑剛 (京大) 岩崎未央 (京大) 河野信 (DBCLS)
五斗進 (DBCLS) 石濱泰 (京大) 瀧川一学 (北大) 吉沢明康 (京大)

目的 質量分析計を用いたプロテオーム解析における精密質量と電荷数の推定

手法 マススペクトルを入力とし、深層学習 (CNN) を用いた学習

結果 既存の手法では取りこぼしていたイオンピークが検出可能となった

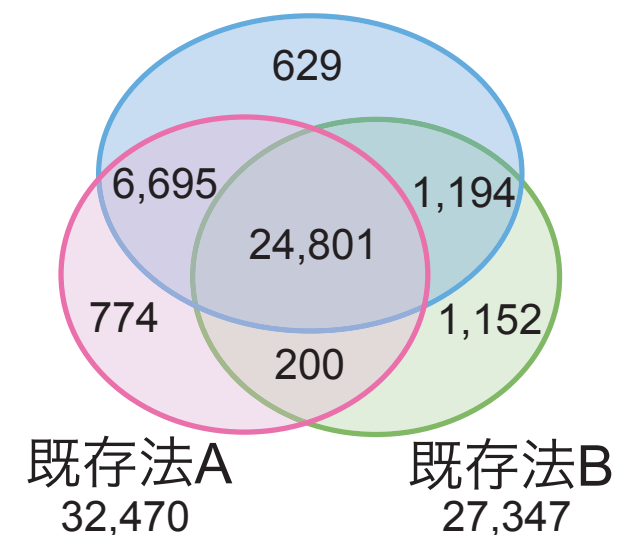
同位体イオンピークが等間隔に現れ、クラスタを形成
ピークの間隔から電荷数を推定可能
クラスタ中で最小 m/z ピークを精密質量のピークとして推定可能



同色で示したピークが同位体ピーククラスタだと考えられる

ピーク検出数

当手法
33,319



知識グラフによる生成モデル学習の正則化

武石 直也 (理研AIP)

目的

知識グラフ等で表される特徴量の事前情報を生成モデルの正則化に活用

手法

decoder and/or encoder の出力を独立性尺度を用いて正則化

結果

汎化性能が改善（予備実験）

- データの特徴量同士の関係を含んだ知識グラフ等 (例: 工場のセンサデータに対して, センサや機器の関係を表したシステム図) を機械学習に利用したい
- 特に, データの生成モデル学習の正則化のために利用することを目指す
- そのために, decoder や encoder にあたる部分の入出力に対して特徴量の(非)類似度をもとに独立性尺度に基づく正則化をかけてみる



低次元サポートをもつ分布の推定について

南 賢太郎（東京大）

目的

密度をもたない分布の推定についての理論構築

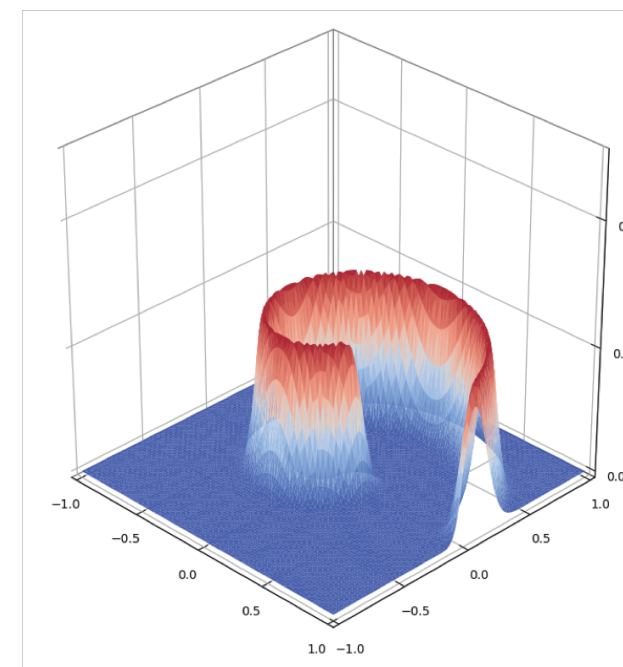
手法

“Entropic Unbalanced Optimal Transport Loss” の最小化

結果

ノイズが畳み込まれた分布に対して，Hellinger距離での収束性を解析

- 画像などの分布のサポートは低次元と言われている
- Wasserstein GANなどの既存手法は低次元性を考慮
- しかし，理論的な正当化はなされていない
- 非対称エントロピー正則化輸送距離を最小化する推定量について考察し，収束レートの解析を行う



リカレントニューラルネットワークにおける ハイパパラメータの最適化と分析

工藤 一輝（東工大） 尾崎 嘉彦（産総研・グリー）
大西 正輝（産総研） 一色 剛（東工大）

目的

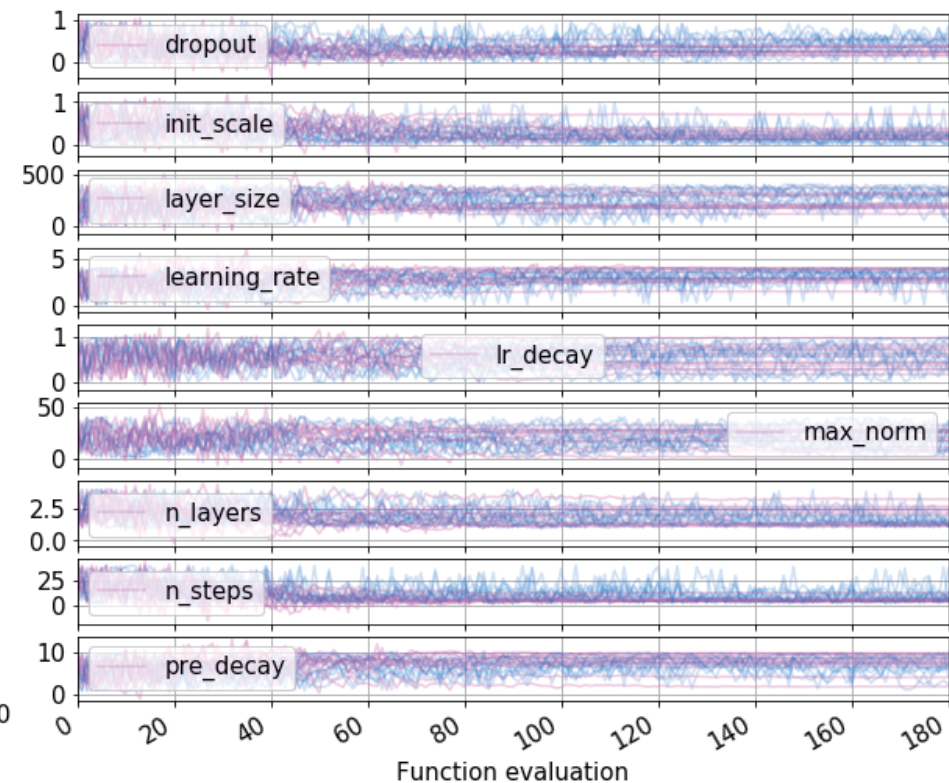
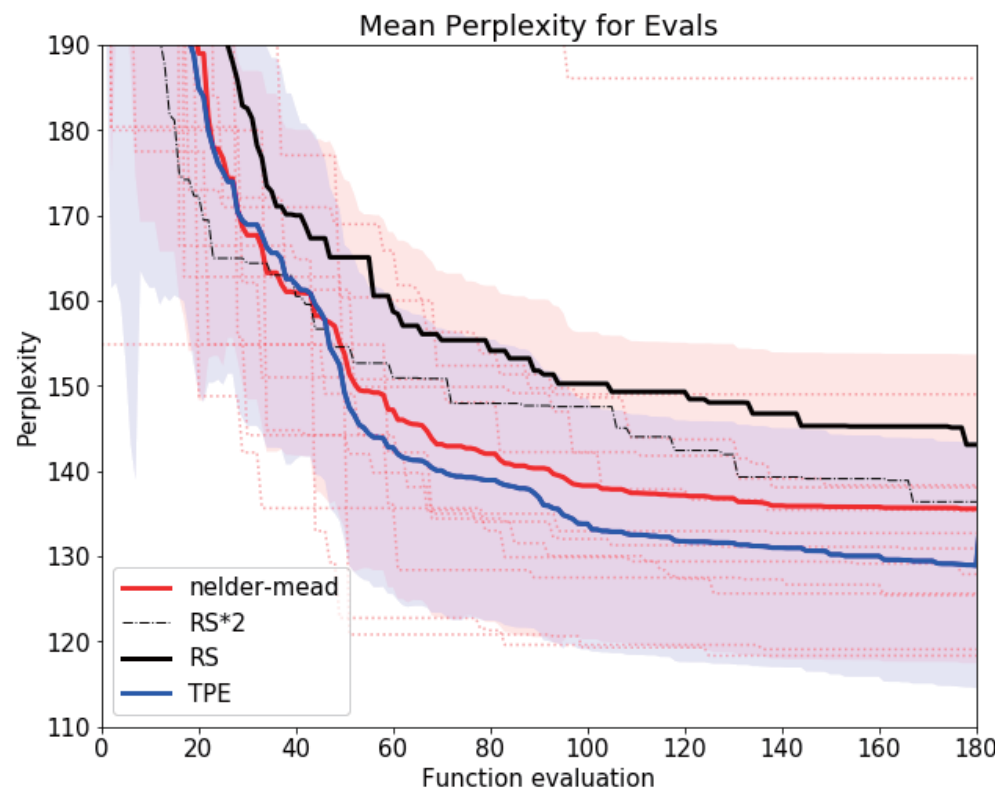
LSTM言語モデルのハイパパラメータ最適化

手法

ブラックボックス最適化手法の適用と比較

結果

Random Search, Nelder-Mead, Tree-structured Parzen Estimatorの比較評価



スパースモデリングの量子多体物理への応用

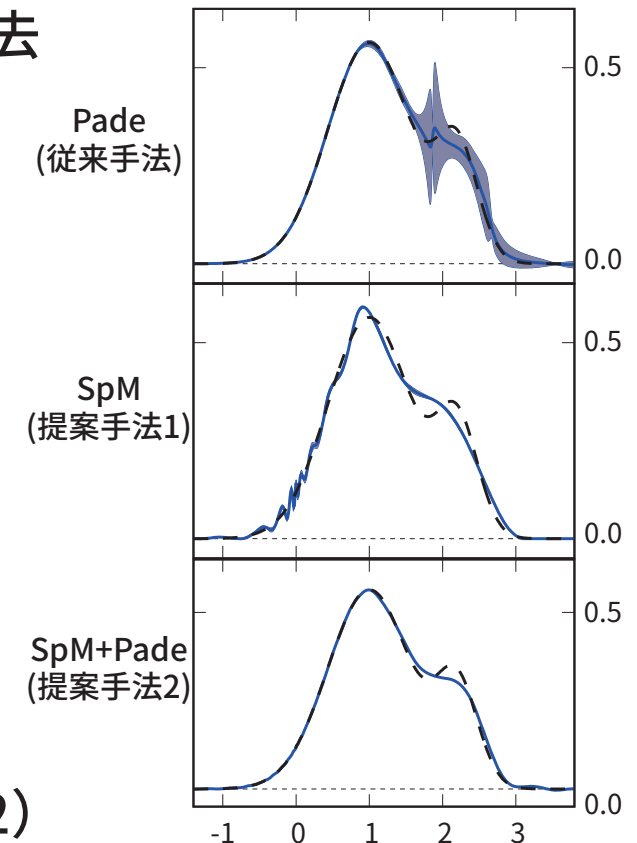
本山 裕一 (東大) 吉見 一慶 (東大) 大槻 純也 (東北大) 品岡 寛 (埼玉大)

目的 ノイズに対して不安定な積分方程式を逆問題として解く

手法 基底変換とスパースモデリングによるノイズ除去

結果 非常に安定な結果を得られた

- スペクトル関数は虚時間グリーン関数から積分方程式の解として得られる (逆問題)
- これは入力のノイズに対して結果が不安定
- 特異値分解を用いた基底変換とスパースモデリングとを組み合わせることでノイズ耐性を得ることに成功した (提案手法1)
- 従来手法と提案手法1 とを組み合わせることで非物理的な振動を抑えることに成功した (提案手法2)



厳密解 (破線) から作った30個のノイズ付きサンプルから各手法で復元した結果 (期待値と標準偏差)

Towards Rank-Loss Optimized Stochastic Block Model

Masafumi Oyamada Koji Ichikawa (NEC)

目的

手法

結果

水道事業における試行錯誤可能なダイナミクス学習

加藤 淳也 (NEC) 比嘉 亮太 (NEC)

目的

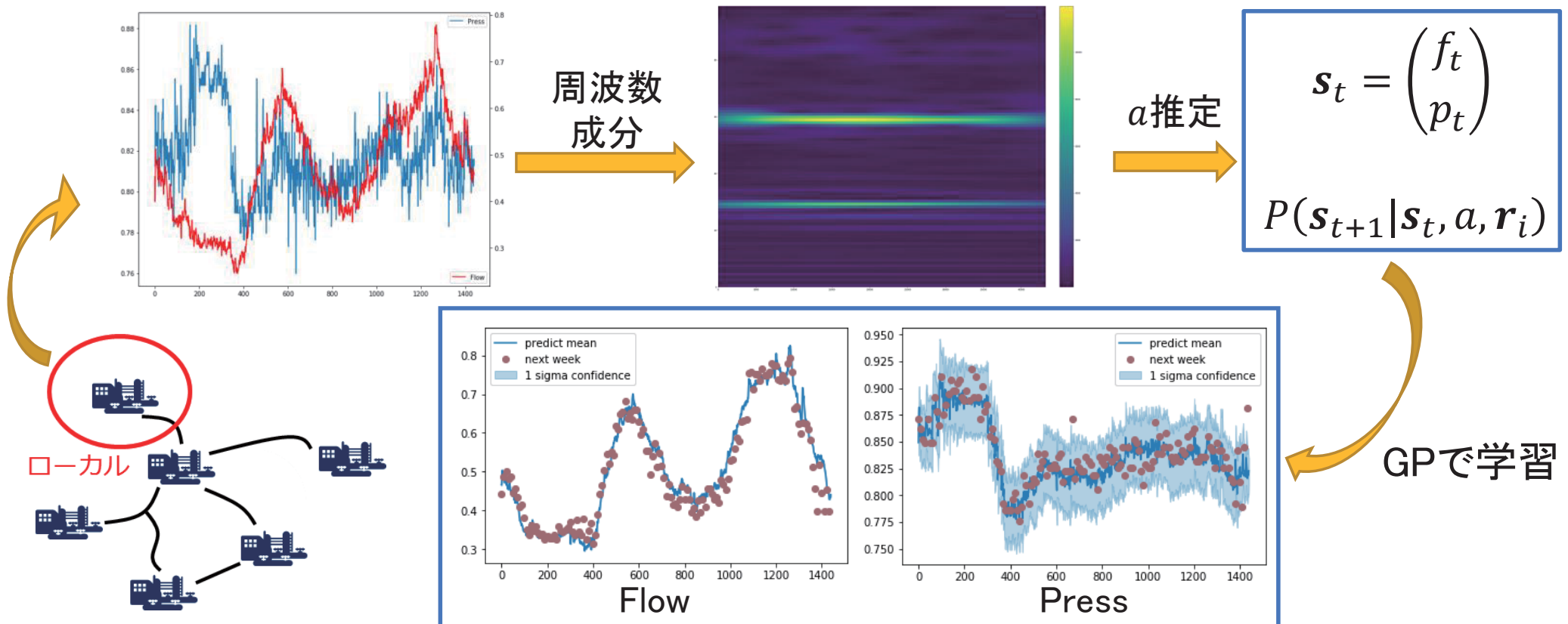
水道事業の運用支援、ダウンサイジング時のシミュレーションの実施

手法

データに含まれる周波数成分からアクションを推定し、GPで個別に学習

結果

一部データでの実験ながら、時系列変化が学習できていることを確認



Learning Temporally Coherent Video Representations using Group Fourier Analysis

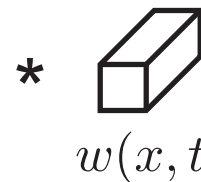
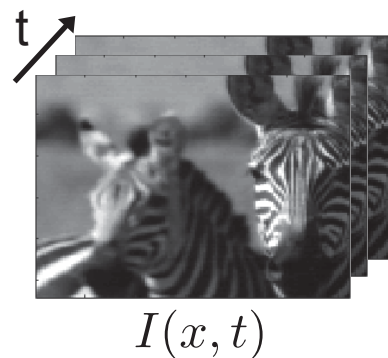
石川 康太 関川 雄介 佐藤 育郎 (デンソーアイティラボラトリ)

目的 時空間信号の物理的に自然な表現学習モデルを構成する

手法 時空間畳み込みフィルタに対し群作用で記述される運動モデルを導入

結果 トーラス群, R の直積群上で線形な運動に対し効率的な学習則を導出

時空間信号に対する畳み込み



フィルタモデル

$w(x, t)$: 3D-CNN [Shuiwang Ji+ 2013]

$w(x, t) = w(x)v(t)$: Decomposed Filter [Lin Sun+ 2015]

$w(x, t) = w(g_{\omega t}^{-1}x)$: Proposed

$g_{\omega t} \in G$: トーラス群, R の直積群 (ω : "速度"パラメータ)

畳み込み公式

$$\int I(x', t') w(g_{\omega(t'-t)}^{-1}(x' - x)) dx' dt' = \sum_l \int \hat{I}(x', -l\omega) \hat{w}^l(x' - x) e^{-il\omega t} dx'$$

I の時間Fourier変換
 ω のG-Fourier変換

Fourier domainで時間・空間依存性を分離することで効率的に畳み込み, および勾配計算

Spectral Parameter Estimation of Random Matrix Models

早瀬 友裕 (東大)

目的

次元削減等に向けた, ランダム行列モデルのパラメータ推定

手法

経験固有値分布にCauchy雑音を加え, ある密度関数の回帰問題に帰着

結果

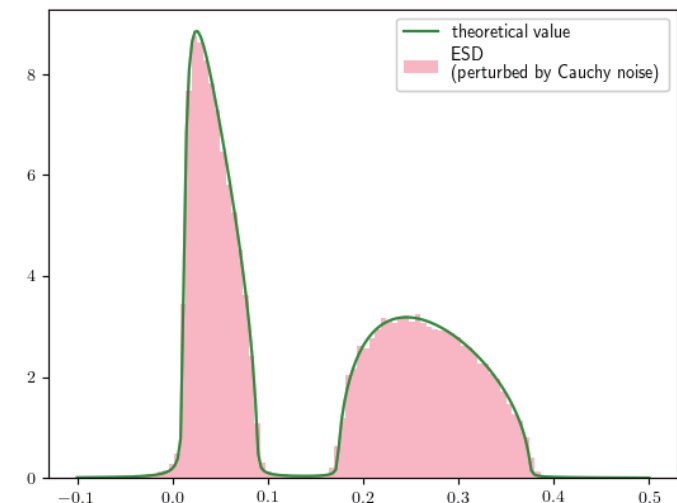
二つのランダム行列モデルのパラメータ推定の実験的成功

- ランダム行列モデルの例:
(Z はGaussianランダム行列)

$$W(D) = Z^* D Z,$$

$$W(A, \sigma) = (A + \sigma Z)^* (A + \sigma Z).$$

- 固有値分布は特異性があるが扱いづらいが,
Cauchy 分布との畳み込みは, 自由確率論による
マスター方程式の解として計算可能
- これとCauchy 雑音を加えた標本固有値分布の
クロスエントロピーの最小化を行う
- 信号が低ランクでなくとも, ランク推定が可能



Spectral Parameter Estimation of Random Matrix Models

Tomohiro Hayase (The University of Tokyo)

目的

手法

結果

人口データからの移動傾向推定の高速化について

赤木 康紀 西村 拓哉 倉島 健 戸田 浩之 (NTT)

目的

時間別・エリア別の人口分布データから移動人数を高速に推定

手法

凸費用輸送問題へ帰着

結果

理論保証付きで高速な推定が可能に！

人口データ からの **移動人数** の推定
([Akagi et al. (IJCAI2018)] など)

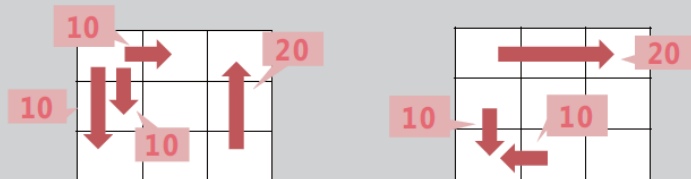
Input: Population of Each Grid Cell at Each Time

7:00 am	8:00 am	9:00 am
90 60 40	60 70 60	40 10 80
10 20 40	20 20 40	10 20 20
30 40 80	40 40 60	60 30 20

Output: # of People Who Moved Between Cells

Between 7:00 am and 8:00 am

Between 8:00 am and 9:00 am



アルゴリズム中で以下の問題を解く必要

$$\begin{aligned}
 &\underset{M}{\text{maximize}} && \prod_{t=0}^{T-2} \prod_{i=0}^{V-1} \left(\frac{N_{ti}!}{\prod_{j=0}^{V-1} M_{tij}!} \prod_{j=0}^{V-1} \theta_{ij}^{M_{tij}} \right) \\
 &\text{subject to} && \underline{N_{ti}} = \sum_{j \in \Gamma_i} \underline{M_{tij}} \quad (t = 0, 1, \dots, T-2), \\
 &&& \underline{N_{t+1,i}} = \sum_{j \in \Gamma_i} \underline{M_{tji}} \quad (t = 0, 1, \dots, T-2), \\
 &&& \underline{M_{tij}} \in \mathbb{Z}_{\geq 0}
 \end{aligned}$$

— : 人口データ
— : 移動人数

既存手法: Stirling の近似・連続緩和 → L-BFGS-B 法

提案手法: 凸費用輸送問題に帰着 → 最短路反復

バイクラスタリングに基づいた がん進化のモデリングと体細胞変異検出への応用

森山 卓也 山口 類 井元 清哉 宮野 悟 (東京大学医科学研究所)

目的 がん組織の進化過程のモデル構築と体細胞変異検出の高精度化への応用

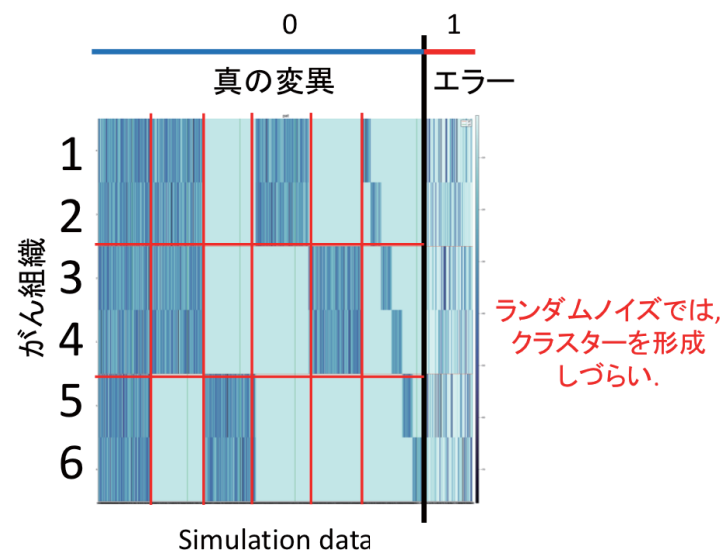
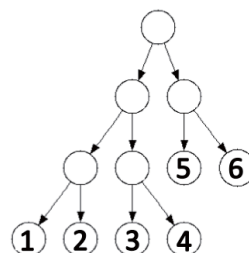
手法 Subset Infinite Relational Models により、がんの進化過程をモデル化

結果 シミュレーションデータを用い、変異検出精度が改善されることを確認

- 変異パターンの導入と性質
 - 各変異に対し、 n 組織のがんのどれが変異を持つかを表す 2 進数を変異パターンとする。
 - 系統樹上では、 $O(n)$ 種類。
ランダムノイズでは $O(2^n)$ 種類。
- 変異パターンの総数の違いから、Subset Infinite Relational Models によるバイクラスタリングに基づいたモデル化を行った。

系統樹上の 変異パターン	ランダムノイズ での変異パターン
$O(n)$	$O(2^n)$

変異パターンに限りがあり、
クラスターを形成しやすい。



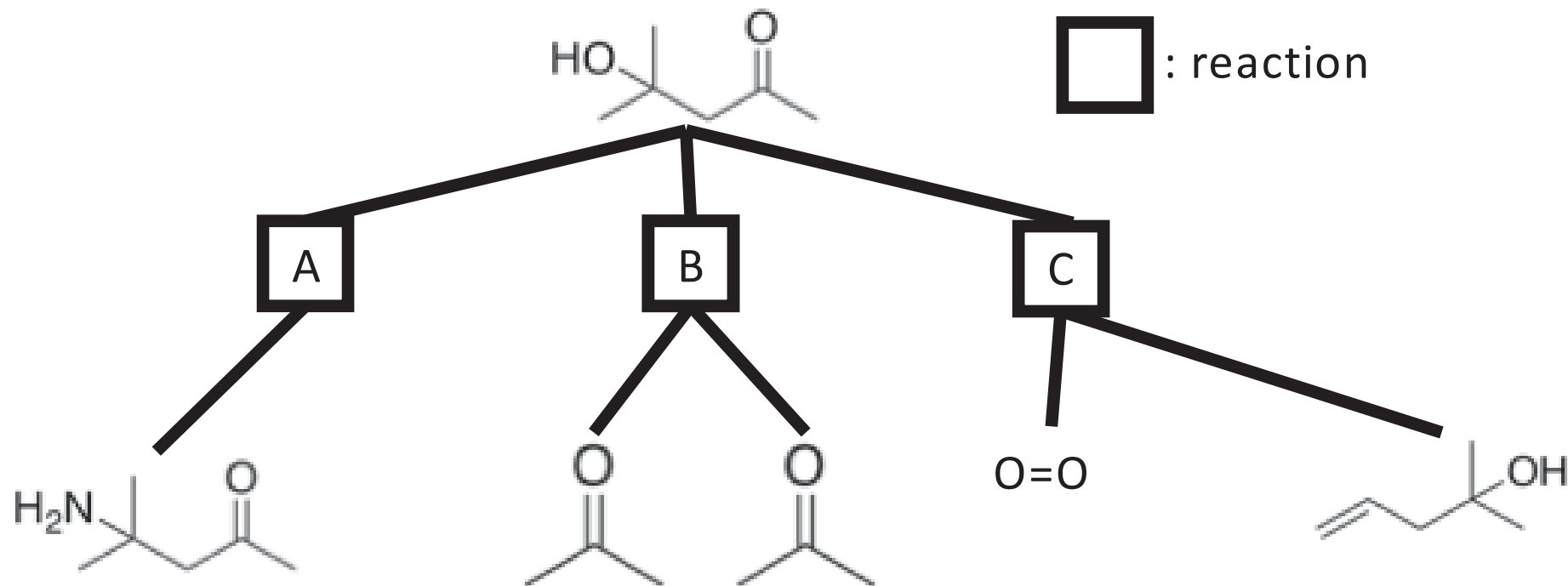
証明数探索を用いた化合物合成経路の列挙アルゴリズム

渋谷 亮祐 石田 祥一 寺山 慧 美添 一樹 奥野 恭史 津田 宏治

目的 化合物の合成経路を列挙する。

手法 AND-OR木による表現に対して証明数探索を応用した。

結果 証明数探索を応用した列挙アルゴリズムを開発した。



患者情報の季節性と検索クエリに基づくトレンドを用いた インフルエンザ患者数予測

村山太一; 若宮翔子; 荒牧英治 (NAIST)

目的

インフルエンザ罹患者数の予測

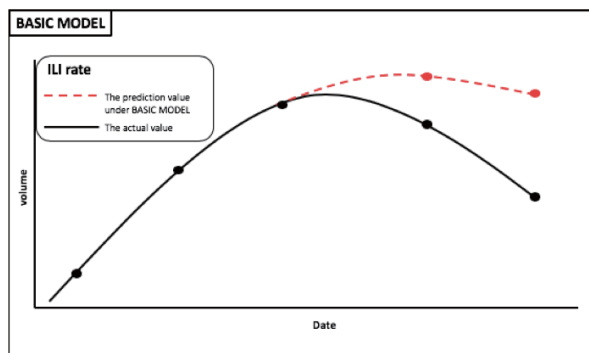
手法

時系列固有部分の予測とトレンド部分の予測の2つのモデルに分割

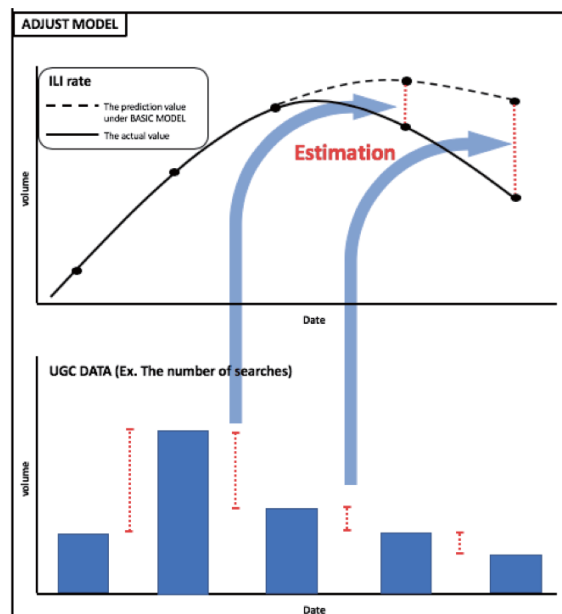
結果

ノイズに頑健性のあるモデルとして、比較手法より高い精度で予測可能

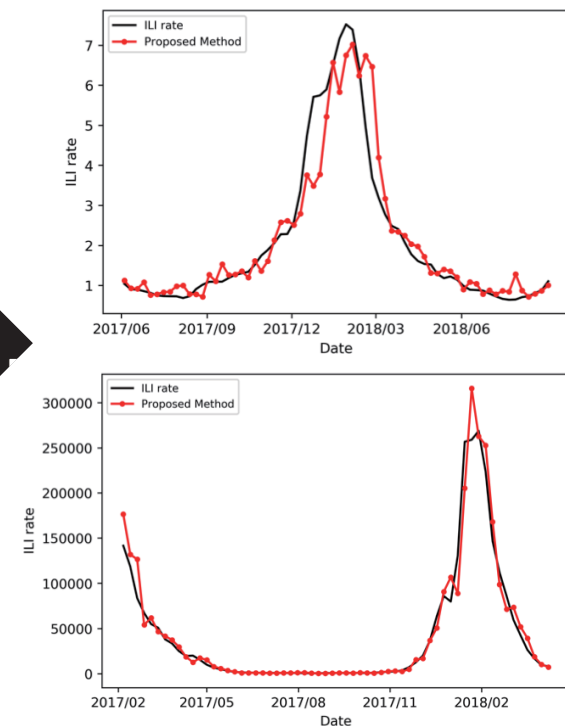
インフル情報からのインフル予測
(時系列固有部分)



Google Trendsを用いて
予測を修正 (トレンド部分)



アメリカと日本のインフル予測



GANの統計的決定理論的定式化と安定化高速化手法

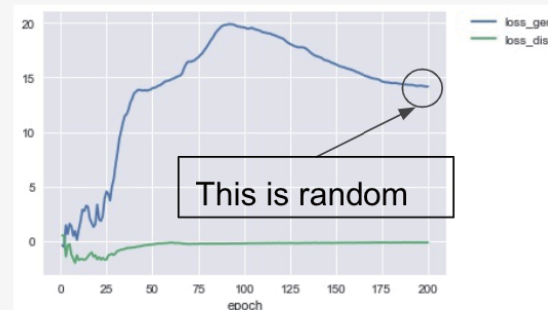
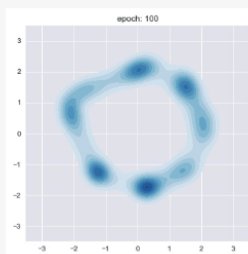
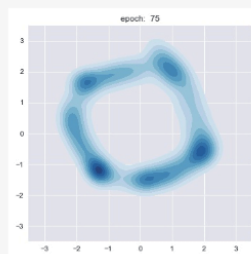
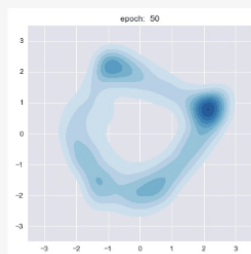
Kei Ikegami (UT-Econ), Kohei Hayashi (PFN), Masaki Watanabe (PFN)

目的 WGANを安定化、高速化させる

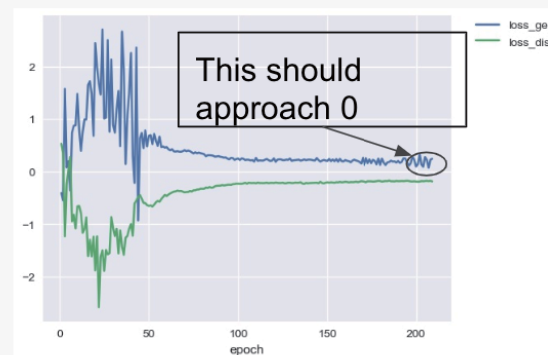
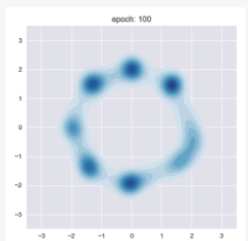
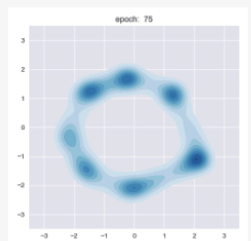
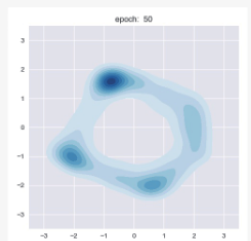
手法 WGANの不決定性を制約付き最適化問題で無くす

結果

normal WGAN_gp : lam 3, epoch 50, 75, 100



Accelerated (and Stabilized)
WGAN_gp : lam 3, epoch 50, 75, 100



表現論を用いた指数型分布族の構成法

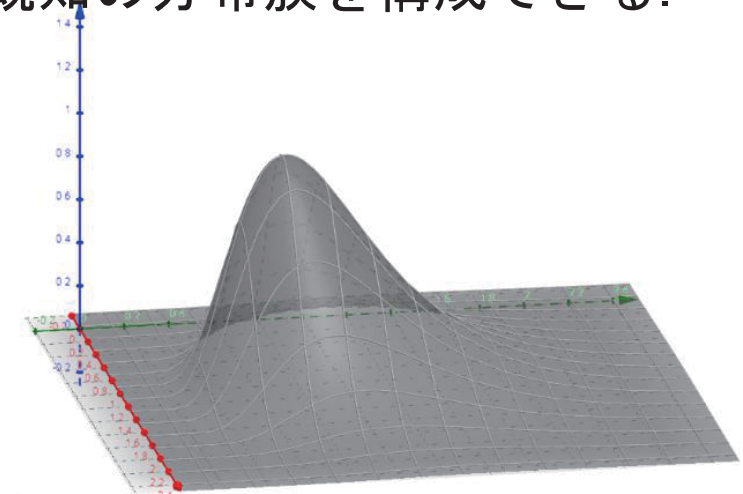
東條 広一 (東京大学) 吉野 太郎 (東京大学)

目的 等質空間 G/H 上に良い指数型分布族を構成する

手法 G の表現から G/H 上の指数型分布族を構成する統一的な手法を導入

結果 既知の分布族に加えて上半平面上の新しい分布族も構成できる

- 適切にリー群 G とその表現を選ぶことで、次の既知の分布族を構成できる。
 - ベルヌーイ分布, カテゴリカル分布,
 - 正規分布, 多変量正規分布,
 - ガンマ分布, 逆ガンマ分布, Wishart 分布,
 - von Mises 分布, Fisher-Bingham 分布.
- 新しい分布族として
 - 上半平面上のポアンカレ計量と相性の良い分布族を構成できる.
- 原理的には限られた分布族しか作れないが多くの有名な分布族を含む.
- 得られる分布族は自然な G 作用で不変である.



Itemsetカーネルに対するランダム特徴

新 恭兵 (北大) Subhransu Maji (UMass Amherst) 小山 聡 (北大)

目的

Itemsetカーネルを用いた際のカーネル法のscalabilityの問題を解決したい

手法

Itemsetカーネルに対するランダム特徴を提案

結果

性能を大きく落とすことなく，省メモリ・高速化を実現

Itemsetカーネル：Itemset集合 $\mathcal{S}$ に含まれる特徴の組合せを扱うカーネル：

$$K(\mathbf{x}, \mathbf{y}; \mathcal{S}) = \sum_{S \in \mathcal{S}} \prod_{j \in S} x_j y_j$$

理論的成果

- Itemset集合に依存しないバウンドの導出
- minimax最適な分散を達成する分布の導出

提案法を用いた線形モデルの利点

- カーネル法では動かない**大規模なデータで動かすことができる**
- Factorization Machineと異なり**最適化問題が凸最適化で扱いやすい**

深層ニューラルネットワークにおける Fisher情報行列の普遍性

唐木田 亮 (産総研) 赤穂 昭太郎 (産総研) 甘利 俊一 (理研CBS)

目的

深層ニューラルネットモデルのパラメータ空間構造を解明

手法

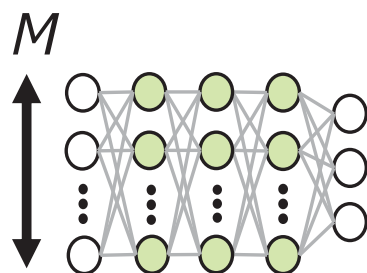
Fisher情報行列の固有値を平均場理論を使って解析的に評価

結果

多数の固有値 (空間の歪みに対応) がゼロに近く, 少数が非常に大きい

深層モデル

- ・ ランダム結合
- ・ 幅 $M \gg 1$ の極限



Fisher情報行列の固有値分布は,

平均 α_1/M 分散 α_2
最大固有値 $\alpha_{max}M$

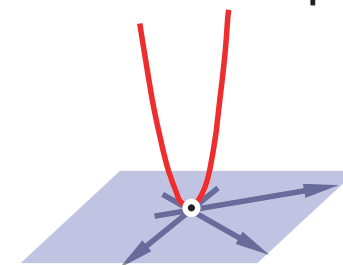
定数 $(\alpha_1, \alpha_2, \alpha_{max})$
は平均場理論から
求値.

様々な層数・非線形性で普遍的に成立

結果の意味:

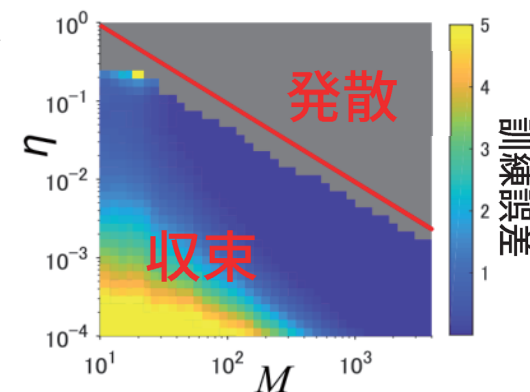
パラメータ空間は
多数の次元が**ほぼ平坦**,
少数が大きく**歪んだ形**

ex. 大域解まわりの
loss landscape



勾配法の発散を防ぐ
学習率 η を
最大固有値 から
推測できる

$$\theta \leftarrow \theta - \eta \nabla_{\theta} L(\theta)$$



Improved mistake bounds for binary matrix completion

Yaxiong Liu (Kyushu U), Ken-ichiro Moridomi (So-net Media Networks),
Kohei Hatano (Kyushu U/AIP), Eiji Takimoto (Kyushu U)

目的

Minimize #mistakes in online binary matrix completion

手法

Reduction to online SDP problem + FTRL with log-det regularizer

結果

Mistake bound i.t.o. hinge loss of **best** linear classifiers over **best** feature space

For each round $t = 1, 2, \dots, T$



Algorithm

1. Receive user-item pair $(i_t, j_t) \in [m] \times [n]$

2. Predict $\hat{y}_t \in \{+1, -1\}$

3. Receive true evaluation $y_t \in \{+1, -1\}$

					
	+		-		-
	-		+	-	
		+		?	-

Theorem

$$\# \text{mistakes} = O \left(\frac{m+n}{\gamma^2} + \inf_{PQ^T \in \mathbb{R}^{m \times n}} \sum_{t=1}^T \text{hinge}_{\gamma}(y_t \bar{P}_{i_t} \bar{Q}_{j_t}^T) \right)$$

- improvement from Herbster et al. 2016
- optimal if separable

Implication: The offline version performs as well as SVMs with the **best** kernel!

One-Class Collaborative Filtering based on Contents Similarity

市川 幸史 (NEC)

目的 Implicit dataに対するCollaborative Filteringの高度化

手法 目的関数にユーザ行動とアイテム関連度を使用した重みを取り入れる

結果 計算量をそれほど大きくせず、従来手法に比べ優れた推薦精度を達成

- 目的関数

$$L = \sum_{u,i} w_{ui} \left(Y_{ui} - \sum_d u_{ud} v_{id} \right)^2 + \lambda \left(\sum_{u,d} u_{ud}^2 + \sum_{i,d} v_{id}^2 \right)$$

- $Y_{ui} = 0$ に対する重み w_{ui} :

- w_{ui} が大きい $\rightarrow$ ユーザ u が商品 i を嫌っている確度が高い。
- w_{ui} が小さい $\rightarrow$ ユーザ u が商品 i を嫌っている確度が低い。

- 発表内容

- ユーザ行動とアイテム関連度を取り入れた w_{ui} の構成方法・精度検証
- 計算量を削減するアルゴリズム

学習曲線の予測に基づく ハイパーパラメータの最適化法

北村 寛 (東大) 前田 新一 (Preferred Networks) 佐野 正太郎 (Preferred Networks)

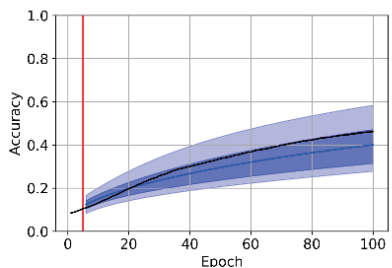
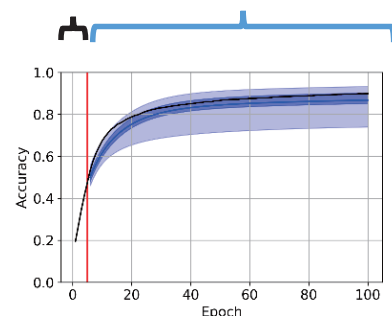
目的 ニューラルネットワークのハイパーパラメータの効率的な最適化

手法 学習曲線の予測を行い有望な候補に対してのみ学習を進める

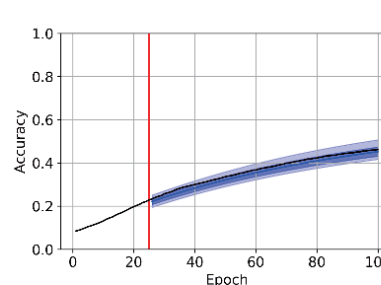
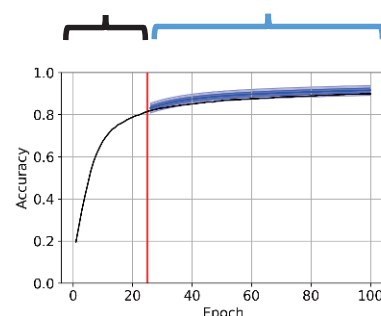
結果 既存手法 Hyperband に並ぶ性能を達成

Bayesian NN による学習曲線予測

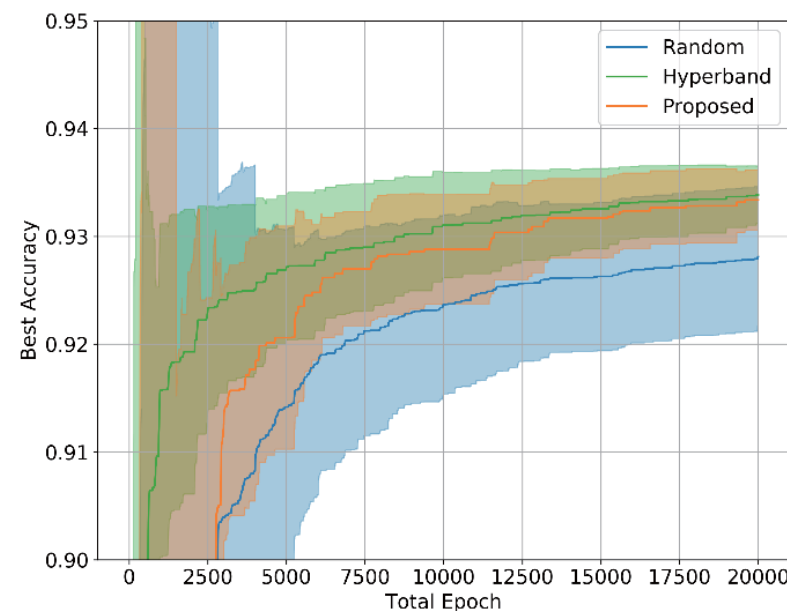
学習済 予測



学習済 予測



ハイパーパラメータ最適化の結果



マルチモーダルカテゴリ形成とSLAMの統合モデルによる オンライン学習と経路計画

谷口 彰 (立命館大/学振PD) 萩原 良信 (立命館大) 谷口 忠大 (立命館大)
稲邑 哲也 (NII/総研大)

目的

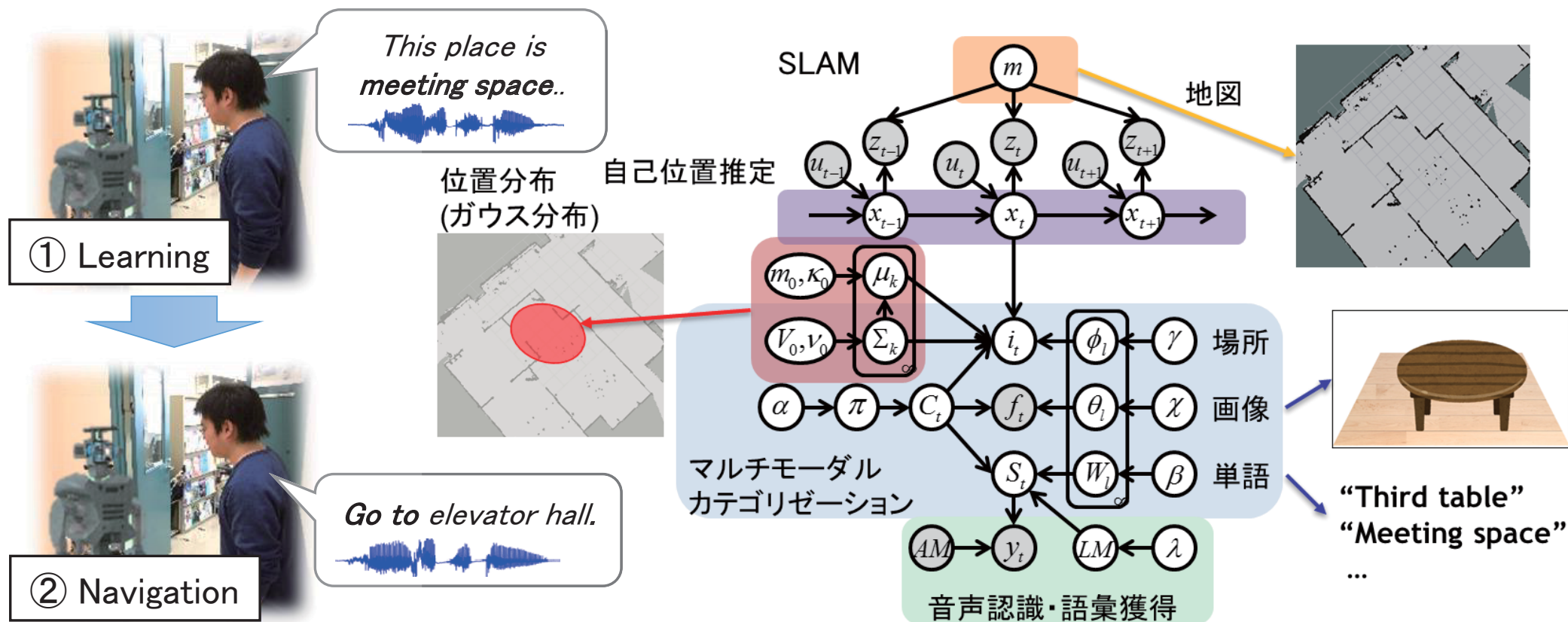
① 確率的生成モデルの教師なしオンライン学習 ② 音声命令からの経路計画

手法

① Particle Filterによる同時事後分布推定 ② POMDPに基づく軌跡の確率推論

結果

① 実環境データを用いた学習を実現 ② 経路計画の定式化



Risk-Aware Reinforcement Learning

綿引隼人 (東大) 前田 新一 (PFN) 岡田 慎太郎 (PFN)

目的

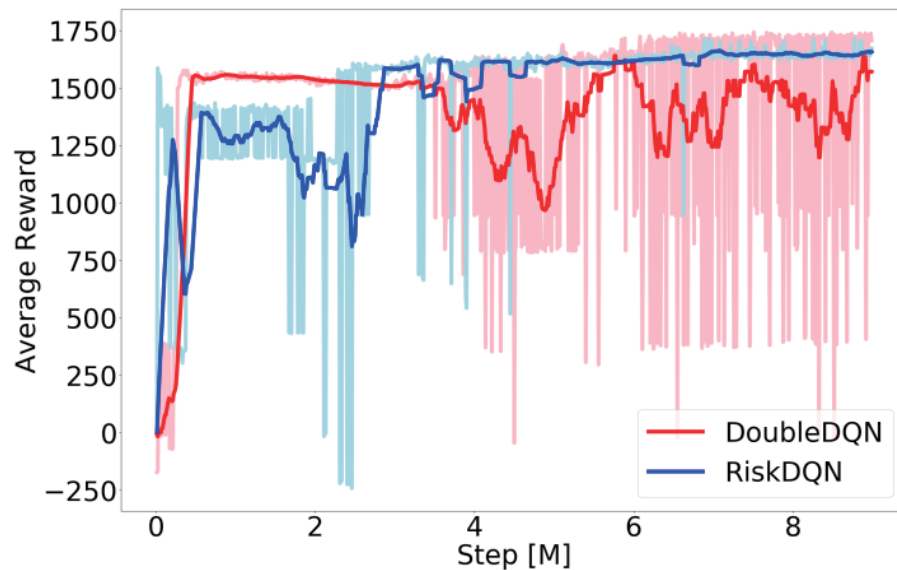
強化学習を訓練中を含め安全に実行する

手法

事前に推定したリスク関数により行動の選択肢を安全なものに制限

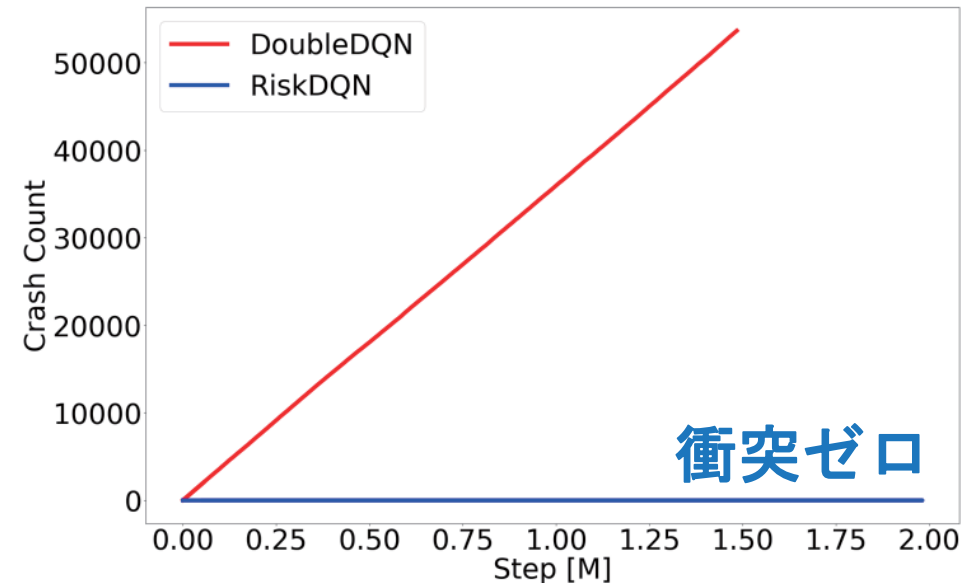
結果

危険な状態を回避しつつ安定した学習が可能



報酬の推移

安定した学習が行えている



累積衝突回数の推移

RiskDQNは衝突せずに学習を完了

Distributionally Robust Learning from Incomplete Data

Amir Najafi (Sharif University of Technology), Shin-ichi Maeda (PFN), Takeru Miyato (PFN)

目的

To extend DRL to semi-supervised learning and analyze its statistical properties

手法

Novel analytical formulation based on new complexity measure

結果

Generalization bound with finite sample size N

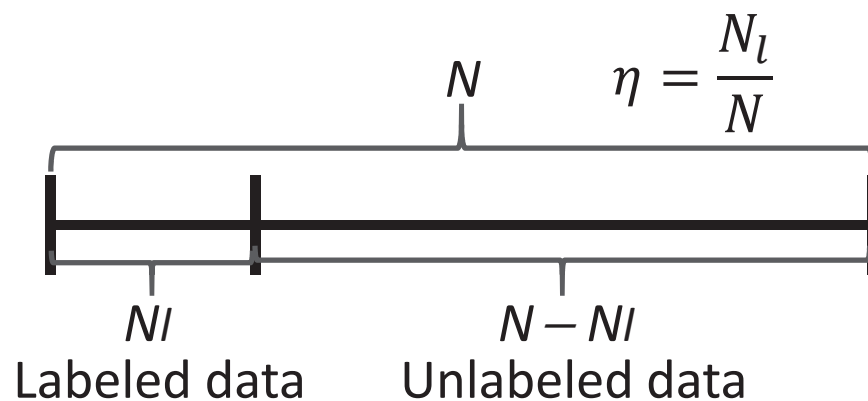
Theorem

With probability at least $1 - \delta$ and under appropriate set of conditions, we have

$$\text{Generalization Error} \leq \text{Empirical Error} + C_{N,\delta} + R_{N,\eta}$$

$C_{N,\delta}$: A constant that decreases with N

$R_{N,\eta}$: A newly introduced complexity measure of loss that decreases with N



Consider the case in which there is no guarantee that the loss function represents the log-likelihood of the true distribution.

- How can we utilize the unlabeled data?
- Is the unlabeled data really useful?

We will answer the questions above!

異質な集団の評価基準差を示すタスクの検出

大橋 英明 (京大) 清水 敏之 (京大) 吉川 正俊 (京大)

目的

集団ごとに評価基準が異なるタスクの検出とそれを考慮した品質管理

手法

潜在変数モデルに基づく推定とタスクの検出を同時に行う手法

結果

実データを用いて検出されたタスクの定性評価を行い考察した

- 集団間の評価基準が異なる例：TAと学生がレポート（タスク）の採点（評価）を行う際、学生はTAと比べフォーマット等の表面的な情報に左右されがち [1]
- 一つのレポートに対する複数の採点結果を集約するとき、集団間の点数の分布が大きく異なるならば集約された値の信頼度は低いと考えられる



- 既存研究：単一の評価基準に基づく値（真の値）の存在を仮定し、評価者ごとの誤差に基づき採点結果が生成されるというモデル [2] ($\odot \rightarrow \otimes \leftarrow \ominus$)
- 本研究：集団ごとに評価基準が異なるタスクの存在を考慮した

[1] Xi Alice Gao et al. Incentivizing Evaluation via Limited Access to Ground Truth: Peer-Prediction Makes Things Worse, In *EC workshop*, 2016.

[2] Panagiotis G. Ipeirotis et al. Quality Management on Amazon Mechanical Turk, In *HCOMP*, 2010.

敵対的学習を用いた知識アラインメント

黒川茂莉^{1,2}、小林亜令¹、萩原将文²
(¹ KDDI総合研究所、² 慶應義塾大学)

目的 一方の知識グラフ(ソース)を用いてもう一方の知識グラフ(ターゲット)を補完

手法 敵対的学習を用いて真の三つ組(正例)と生成する偽の三つ組(偽正例)の同化を図り、ターゲットにとって良質な知識を生成

結果 オープンな知識グラフ(DBP15Kの中国語→英語)の場合で、BootEA(IJCAI18)に対し学習初期の精度上昇を確認
(追加実験中)

- 前提: 一部の实体アラインメントが所与
- 学習方法

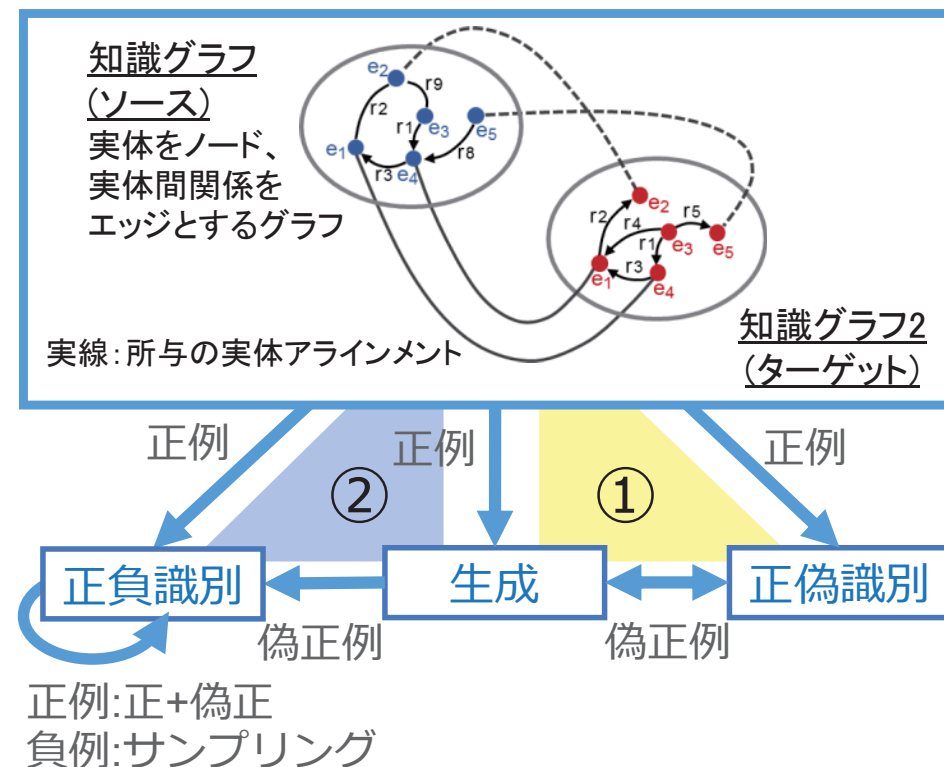
【ステップ①】

敵対的学習を用いて三つ組(偽正例)生成器を学習
※学習方法はKBGAN(NAACL-HLT 2018)と同様

【ステップ②】

生成した三つ組(偽正例)を正例と混ぜて識別器を作成
※学習方法は偽正例の作り方以外はBootEAと同様

- 利用方法
- 上記ステップで獲得した実体、関係のベクトル表現を使って隠れた三つ組を推論



日時が曖昧に記録されたイベント列の解析

勝木 孝行 (IBM) 恐神 貴行 (IBM) 古関 聰 (IBM) 大野 正樹 (IBM) 工藤 道治 (IBM)
牧野 真樹 (藤田医科大学) 鈴木 敦詞 (藤田医科大学)

目的

観測者の即時的な制御無しに受動的に観測されるイベント列の解析

手法

動的ボルツマンマシンを拡張したモデルを提案

結果

実データを用いた実験において有効性を確認

- 本発表では観測者の即時的な制御無しに受動的に観測されるイベント列の解析を扱う。この場合、記録されたタイムスタンプは曖昧であり、イベントが発生した日時を表すのではなく、単にそれが記録・観測された日時を表す。時間間隔は一定ではなく、タイムスタンプの不確実性も大きい。このモデル化に動的ボルツマンマシンを拡張する形で導出される時間割引畳み込み構造と動的プーリングを提案し、実データへの有効性を確認する。

謝辞

本研究の実施にあたり支援をして頂いた第一生命保険株式会社拝田恭一様に深謝いたします。本研究は、JST、CREST、JPMJCR1304の支援を受けたものです。

Variational Inference for Mixture of Finite Mixtures with Unknown Number of Components

JiYao Li (Osaka University)

Yutaka Kano (Osaka University)

目的

To find a variational inference algorithm for Mixture of Finite Mixtures (MFM).

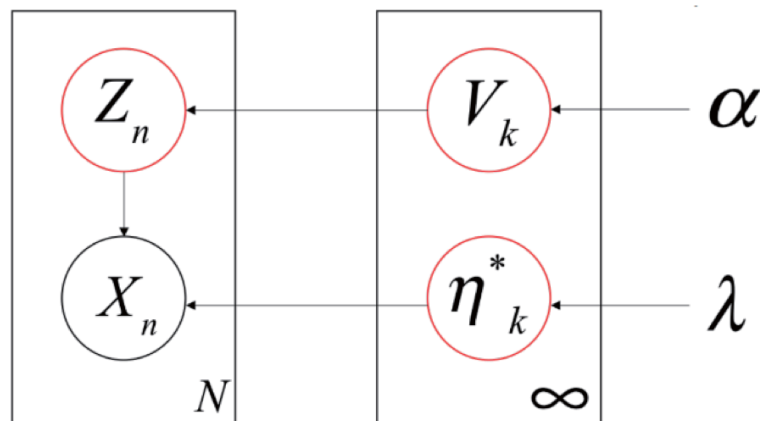
手法

Reformulate the latent variable in MFM using stick-breaking representation, and then apply the mean-field approximation.

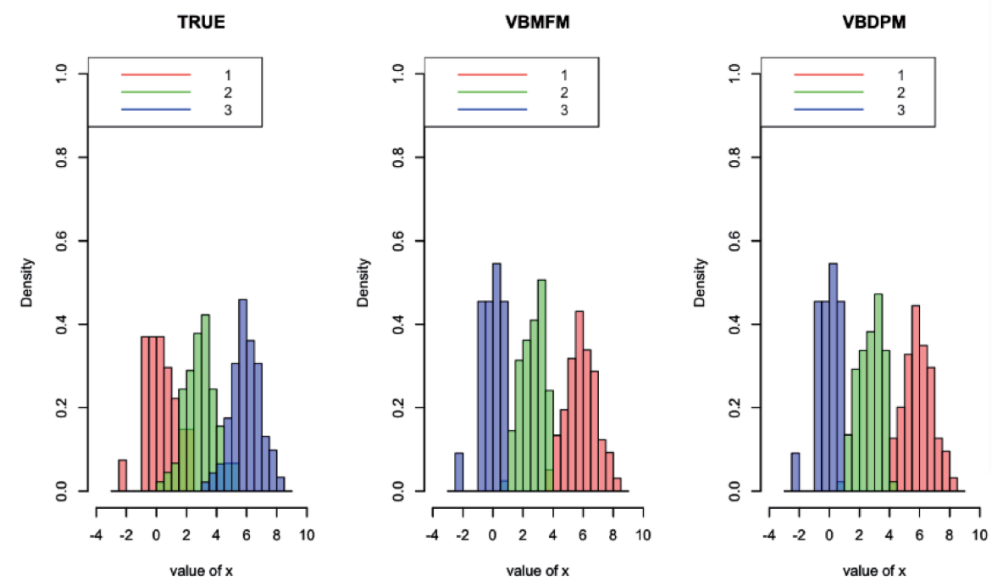
結果

Propose a high-speed clustering algorithm for MFM which is inspired by Dirichlet Process Mixture Model (DPM).

- Graphical representation for exponential family MFM:



- Simulation result:



敵対的複数選択バンディット問題のミニマックスリグレット

伊藤 伸志(東大) 波多野 大督(理研AIP) 澄田 範奈(首都大学東京) 竹村 慧(NEC)
福永 拓郎(理研AIP, JSTさきがけ) 垣村 尚徳(慶應義塾大学) 河原林 健一(NII)

目的

複数選択バンディット問題の最悪ケース解析

手法

非独立な確率分布に従う入力分布の構成

結果

リグレット下界の改善

[問題設定]

入力: 報酬ベクトル $r_t \in [0,1]^d$

For $t = 1, 2, \dots, T$

- アクション $a_t \in A$ を選択
($A \subseteq \{a \in \{0,1\}^d \mid \|a\|_0 = k\}$)
- 報酬 $r_t^\top a_t$ を観測

評価指標: リグレット R_T

$$R_T := \max_{a \in A} \sum_{t=1}^T r_t^\top a - \sum_{t=1}^T r_t^\top a_t$$



リグレット上界 (アルゴリズムの性能)

[Bubeck et al. (2012)],
[Audibert et al. (2013)]:

$$R_T = O\left(\sqrt{k^3 d T \log \frac{ed}{k}}\right)$$

リグレット下界 (最悪ケース解析)

[Cohen et al. (2017)]:

$$R_T = \Omega(\sqrt{k^3 d T / \log T})$$

[本研究]:

$$R_T = \Omega(\sqrt{k^3 d T})$$

組合せ線形バンディットに対するトンプソンサンプリング

竹村 慧(NEC) 伊藤 伸志(NEC)

目的

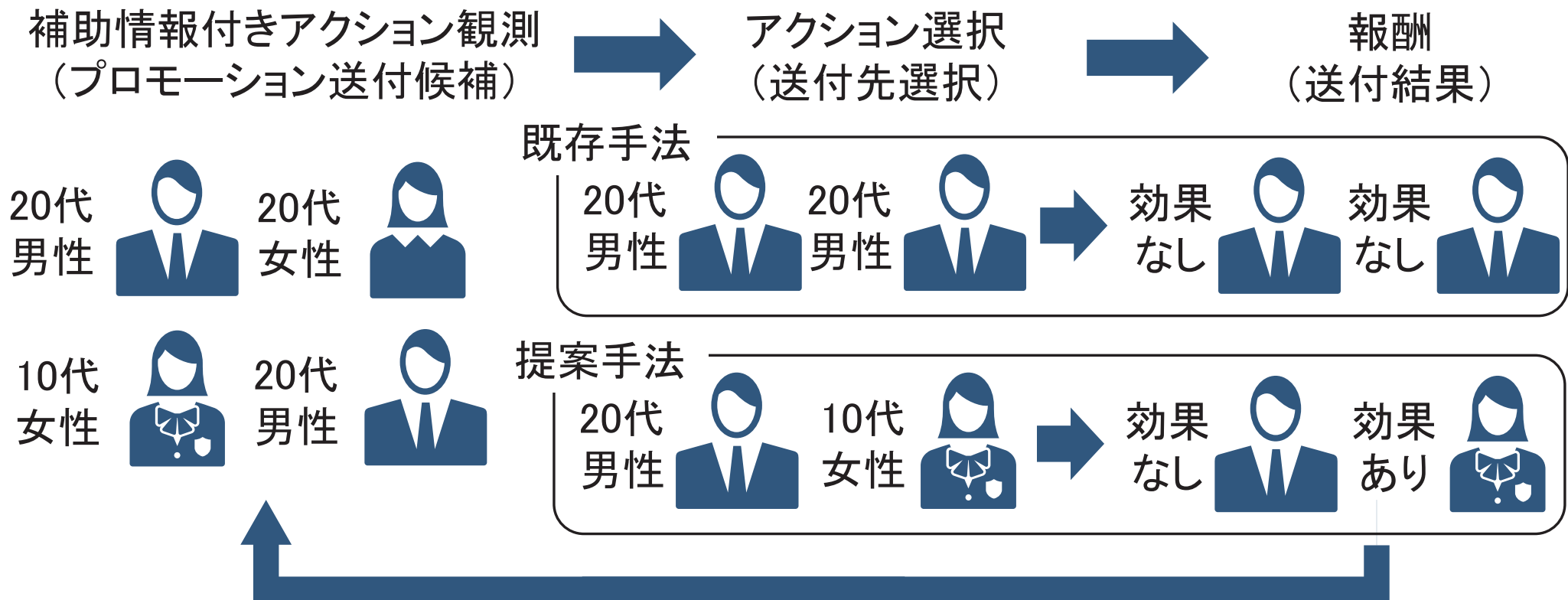
組合せ線形バンディットに対する効率的なアルゴリズムの提案

手法

アクションごとに独立にパラメータをサンプリングして報酬を推定

結果

提案手法は異なる特徴をもつアクションを一度に選べるため効率よく学習



Adaptive Sample-space & Adaptive Probability Coding

中西 健 (東大) 前田 新一 (PFN) 宮戸 岳 (PFN) 小山 雅典 (PFN)

目的

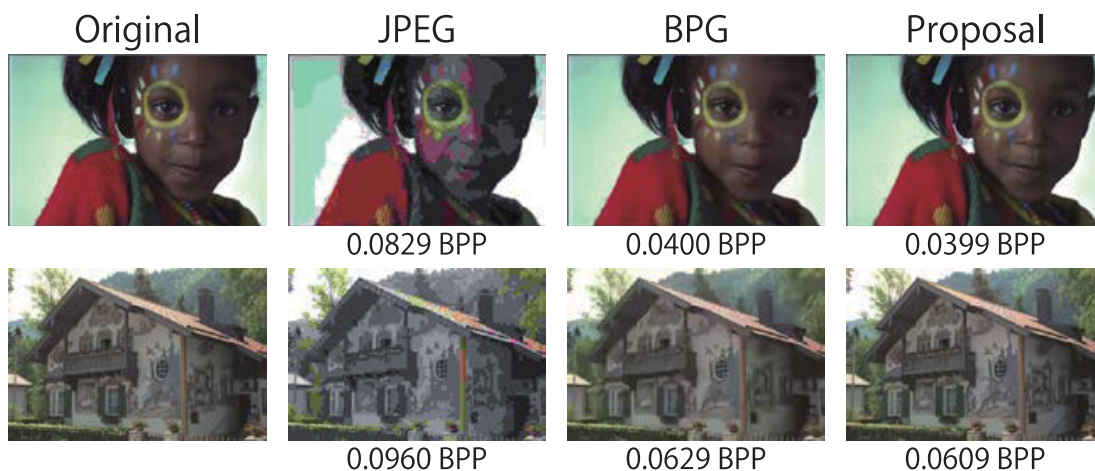
圧縮性能の良い非可逆データ圧縮手法の開発

手法

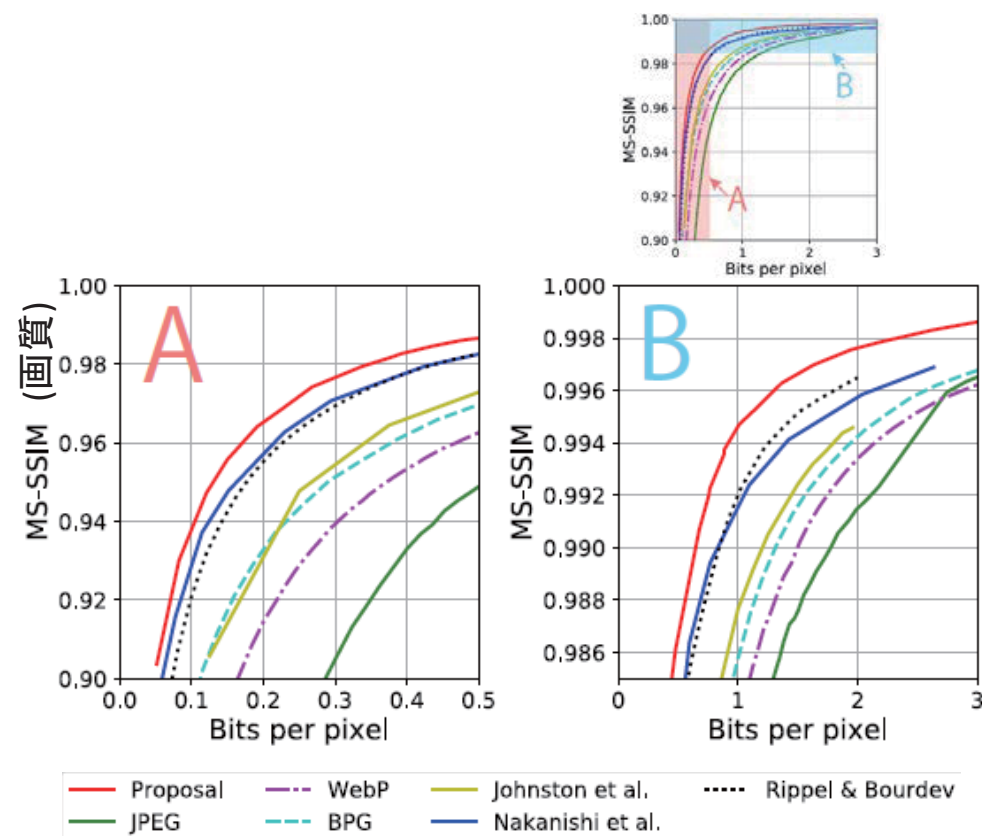
コンテキストに依存して量子化幅を決定するNN量子化・符号化器を提案

結果

state-of-the-artの性能を達成



BPP: Bits per pixel



Relaxing Imbalance with Positiveness

谷本 啓 (NEC) 山田 聡 (NEC) 竹之内 高志 (はこだて未来大) 鹿島 久嗣 (京大)

目的 ニアミスサンプルを活用して不均衡な判別問題をサンプル効率化

手法 一般化蒸留によるソフトラベル化+リバランス

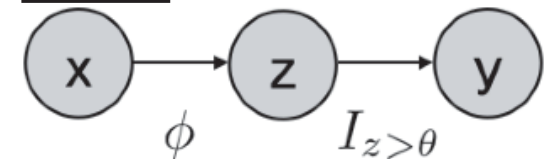
結果 理論的に推定誤差を抑えられ、実験的にも回帰等の既存手法より汎用的

- 不均衡データだが、ラベルが閾値関数によって定義され、ニアミスサンプルが多数得られている想定
- 正負をバランスさせたコスト考慮型損失において有効サンプル数は調和平均となる（線形モデル+ERM）

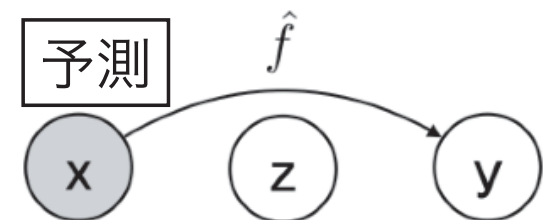
$$\mathbb{E}_S [L_D(\hat{w})] - \inf_{w: \|w\|_2 \leq 1} L_D(w) \leq \mathbb{E}_S \left[\max_n \|x_n\|_2 \sqrt{\frac{1}{N_+} + \frac{1}{N_-}} \right]$$

- 目的変数をソフトラベル $\sigma(z/T)$ としリバランスするとボトルネック $\frac{1}{N_+}$ が緩和
- 実験的にも通常の判別、回帰、ランク学習と比べ、データによらず高精度

学習



予測



適応的な部分グラフ指示子の探索・選択に基づく 非線形グラフ分類回帰

白川 稜 (北大) 横山 侑政 (北大) 岡崎 文哉 (北大) 瀧川 一学 (北大)

目的 グラフに対する教師付き学習（分類・回帰）

手法 部分グラフ指示子の適応的探索・選択に基づく回帰木アンサンブル

結果 部分グラフ特徴集合と非線形な仮説の効率的な同時学習を実現

学習に用いる部分グラフを適応的に探索・選択する回帰木学習アルゴリズムと
勾配ブースティングによるアンサンブル学習を提案する

- 部分グラフパターン探索

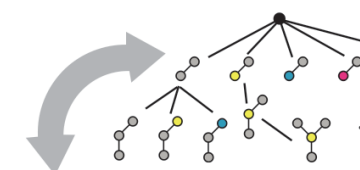
gSpanアルゴリズムによる出現部分グラフの列挙木上の探索

- 変数集合に探索木を伴う場合の回帰木学習

変数集合の木構造に基づく枝刈り設計による変数の探索・選択

- 勾配ブースティングによるアンサンブル学習

精度と不安定性向上のため勾配値近似によるアンサンブル構築



y	G									...
0.1		1	1	1	1	1	1	1	1	...
0.7		1	1	1	0	1	1	1	1	...
0.9		1	1	1	0	1	1	1	1	...
$\vdots$	$\vdots$									...
1.2		1	1	1	0	1	1	1	1	...

Shirakawa R, Yokoyama Y, Okazaki F, Takigawa I.

The 14th International Conference on Mining and Learning with Graphs (MLG 2018)

KDD'18 Workshop, London, U.K., August 20, 2018. <http://www.mlgworkshop.org/2018/>

Wasserstein距離を用いた転移学習の理論解析

熊谷 亘 (理研AIP) 松井 孝太 (理研AIP)

目的 転移学習における期待リスクの評価

手法 Wasserstein距離の双対表現

結果 弱い仮定の下で目標ドメインでの期待リスクの上界を導出

- 結合分布の間の**Wasserstein距離**に対する**数学的な不等式**を導出
- 上記の不等式の**転移学習における意味付け**
- Wasserstein距離を用いた**既存の上界の弱点**を指摘
- 転移学習における最近の**他の理論評価との関係**についても言及

期待リスクの上界に関する詳細はポスターで御覧ください！

概況文作成支援のための予想天気図類似度算出法

雲居 玄道（早大） 後藤 正幸（早大） 吉開 朋弘（日本気象協会）

目的

概況文作成支援のために、概況文の類似度が高いと予想される日を72時間予想天気図より推定する

手法

72時間予想天気図における気圧配置から類似度を算出するため、画像のエッジ検出で用いられるラプラシアンフィルタを導入する

結果

提案した特徴量を用いることで、作成される概況文の類似度が高い日をTopN精度を用いて評価し有効性を示した。

72時間予想天気図：1日後から3日後までの3枚の画像で構成される予想天気図



図2. ラプラシアンフィルタ後

対象データ：2017年1, 4, 7, 10月の123日間
評価指標：TopN精度

該当日の予想天気図に対して、Top10の類似日中に類似度の最も高い概況文の日付が含まれるか否か。

表1.実験結果

特徴量	TopN精度
図1のみ	0.6911
図2のみ	0.6667
図1+2	0.7154

Lipschitz-Margin Training: Scalable Certification of Perturbation Invariance for Deep Neural Networks

都竹雄介 (東大・理研AIP) 佐藤一誠 (東大・理研AIP) 杉山将 (理研AIP・東大)

目的 悪意のあるノイズが加えられた入力に対する堅牢性の保証されたネットワークの構築

手法 ネットワークの勾配の最大値と出力マージンの比の最大化を行う

結果 比較的大きなモデルに対してもノイズ耐性の保証を与えた

予測のノイズ普遍性を保証する十分条件:

$$\underbrace{F(X)_{t_X} - \max_{i \neq t_X} F(X)_i}_{\text{Margin}} > \sqrt{2} \underbrace{L}_{\text{Lipschitz constant}} \underbrace{\|\epsilon\|_2}_{\text{Noise}}$$

$F(X)_i$: ネットワークの出力ベクトルのクラス*i*に対応する要素 t_X : 正解ラベル

クラウドソーシングモデルのミニマックスエラー解析と、 そのワーカークラスタリングモデルへの応用

今村 秀明 (東大・理研AIP) 佐藤 一誠 (東大・理研AIP) 杉山 将 (理研AIP・東大)

目的

正解ラベルが存在しないクラウドソーシングモデルにおいて、適切なモデルの選択が難しい問題を解決する。

手法

Fanoの不等式を用いて、混同行列を用いるクラウドソーシングモデルの理論的な解析を行う。

結果

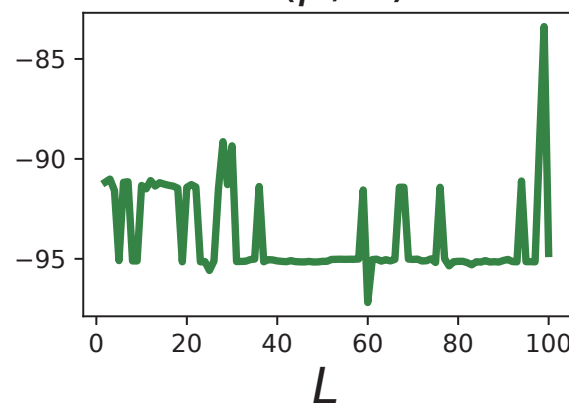
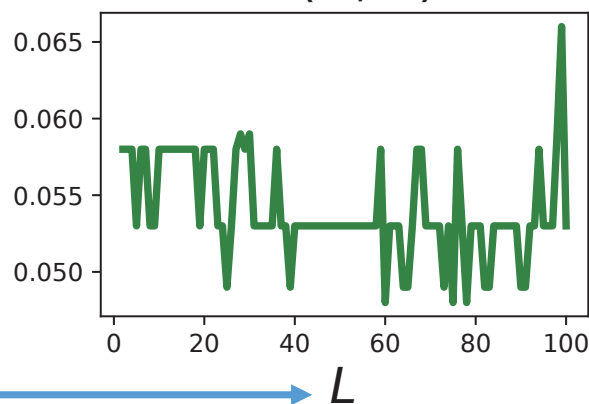
データから計算できる下界を導出し、得られた下界を利用すればモデル選択ができることを実験的に確認した。

$$\inf_{\hat{G}} \sup_{G \in [K]^n} \mathbb{E}[\mathcal{L}(\hat{G}, G)] \geq \frac{1}{n \log K} \left(R(\rho, \pi) - \frac{\log 2}{n} \right)$$

正解ラベル無しには計算不可能な
エラーのミニマックス値 $\mathcal{L}(\hat{G}, G)$

正解ラベル無しに計算可能な左辺の下界
 $R(\rho, \pi)$

モデルを決める
パラメータ



両者は良く似た
振る舞いをする

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

奥野 彰文 (京大・理研AIP) Kim Geewook (京大・理研AIP) 下平英寿 (京大・理研AIP)

目的

1. ニューラルネットワーク(NN)を用いた既存のグラフ埋め込み(GE)の表現能力を理論的に評価し,
2. より高い表現能力を持つグラフ埋め込みを提案する.

手法

1. マーサーの定理とNNの表現定理を組み合わせることでNNを用いた既存のGEの表現定理を示し,
2. 内積類似度(IPS)にシフト項を加えることで既存のGEで表現できない類似度も近似できるようにした.

結果

提案するShifted IPSを用いたGEは既存のGEより表現能力が高いことを示した.

(※ 提案モデルはポアンカレ埋め込み(Nickel and Kiela, NIPS2017)なども近似できる)

観測値: link weights w_{ij} , data vectors x_i

確率モデル: $(w_{ij} \mid \mathbf{x}_i, \mathbf{x}_j) \sim \text{Bernoulli}(\sigma(h(\mathbf{x}_i, \mathbf{x}_j)))$

既存のモデル: Inner Product Similarity (IPS)

$$h(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{f}_{\text{NN}}(\mathbf{x}_i), \mathbf{f}_{\text{NN}}(\mathbf{x}_j) \rangle \text{ (ただし } \mathbf{y}_i = \mathbf{f}_{\text{NN}}(\mathbf{x}_i) \text{ は埋め込み)}$$

提案するモデル: Shifted Inner Product Similarity (SIPS)

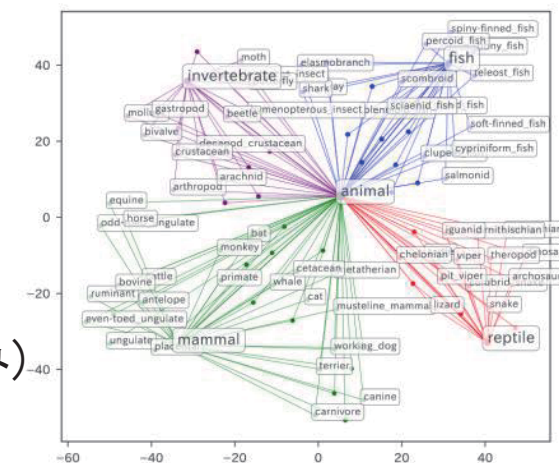
$$h(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{f}_{\text{NN}}(\mathbf{x}_i), \mathbf{f}_{\text{NN}}(\mathbf{x}_j) \rangle + u_{\text{NN}}(\mathbf{x}_i) + u_{\text{NN}}(\mathbf{x}_j)$$


図1. WordNet (Miller, 1995) をSIPSで埋め込んだ結果

IPS $\Rightarrow$ Positive Definite (PD)な類似度のみ近似
SIPS $\Rightarrow$ Conditionally PDな類似度 (Poincare距離,
Wasserstein距離など)も近似できる

詳細は Okuno, Akifumi, Geewook Kim, and Hidetoshi Shimodaira.

"Graph Embedding with Shifted Inner Product Similarity and Its Improved Approximation Capability." *arXiv preprint arXiv:1810.03463* (2018).

単調集合関数制約付き最小化問題に対する貪欲法と射影勾配法

坂上 晋作（NTT コミュニケーション科学基礎研究所）

目的 様々な非ゼロパターン制約を持つ最適化問題に対する手法の解析

手法 貪欲法・射影勾配法

結果 目的関数の制限条件数および制約関数のパラメタ依存の理論保証の証明

$$\text{minimize}_{\mathbf{x} \in \mathbf{R}^{[d]}} l(\mathbf{x}) \quad \text{subject to } G(\text{supp}(\mathbf{x})) \leq c$$

目的関数 l の扱いやすさ（制限条件数）と単調集合関数 G のモジュラ関数への近さ（優加法性比および逆曲率）を用いて貪欲法・射影勾配法の理論保証を示す

応用

連続的スパース性を持つ信号復元

優加法的コスト付き特徴量選択（負荷のかかる検査でデータを得る場合など）

複数観測データによる 熱波の時空間変動エミュレーション

村上 大輔 (統数研) 松井 知子 (統数研) Gareth W. Peters (Heriot-Watt大)

目的 都市熱波状況の空間詳細な把握

手法 Local Tukey g-and-h (local TGH)過程による期待二乗誤差最小化

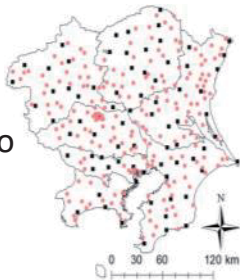
結果 熱波状況の空間詳細な状況を、分布の裾特性も含めて明らかとした

- 衛星観測と地上観測を組み合わせた時空間モデリングによって空間詳細な熱波状況を再現し、Local TGH過程を用いてその背後に潜む局所分布特性を解析

暑熱観測データ

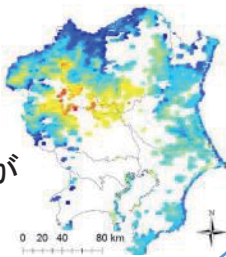
地上観測

- 気象庁+
NTT DoCoMo
- 正確だが、
空間的に疎



衛星観測

- MODIS
- 空間詳細だが
時間的に疎



Local TGH過程

- 地点 s_* の熱波状況 $Y(s_*)$ を周辺 s の観測情報 $Y(s)$ でモデル化

$$\begin{pmatrix} Y(s) \\ Y(s_*) \end{pmatrix} = \begin{pmatrix} a(s) \\ a(s_*) \end{pmatrix} + \begin{pmatrix} b(s) \\ b(s_*) \end{pmatrix} \begin{pmatrix} \tau_{g,h}[Z(s)] \\ \tau_{g,h}[Z(s_*)] \end{pmatrix}$$

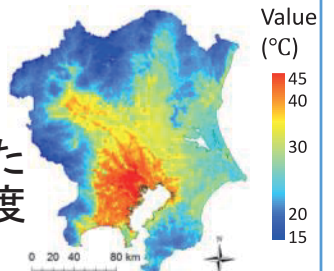
$$\begin{pmatrix} Z(s) \\ Z(s_*) \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} C(s, s) & c^T(s, s_*) \\ c^T(s, s_*) & c(s_*, s_*) \end{bmatrix} \right)$$

$$\tau_{g,h}[Z(s_*)] = \frac{e^{gZ(s_*)} - 1}{g} e^{\frac{hZ(s_*)^2}{2}}$$

g : 歪度を推定
 h : 尖度を推定

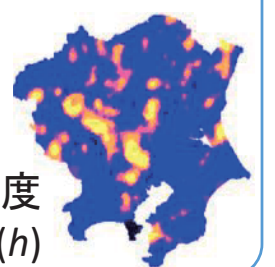
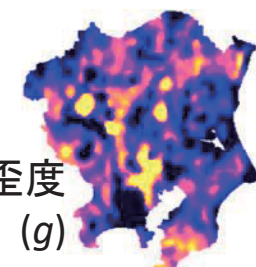
結果(2012/8/19 13:00)

推計された
地表面温度



歪度
(g)

尖度
(h)



都心部を中心とした 住宅地の地価予測アプローチの比較

JAEKEL Marc-Olivier (FPTソフトウェアジャパン株式会社)
中川 帝人 (株式会社オープンハウス)

目的

手法

結果

転移元データのラベルを活用した分布間距離に基づく 教師なしドメイン適応

黒木 誠一 (東大・理研) チャロエンパックディーノントワット (東大) 包含 (東大・理研)
本多 淳也 (東大・理研) 佐藤 一誠 (東大・理研) 杉山 将 (理研・東大)

目的 教師なしドメイン適応

手法 転移元ラベルを活用した分布間距離の提案

結果 汎化誤差上界を持つ, 計算量の小さい分布間距離

	S-disc [提案手法]	X-disc [Mansour et al., 2009]	d_H [Ben-David et al., 2007]
転移元ラベルの活用	○	×	×
汎化誤差上界	○	○	×
計算量オーダー	$\mathcal{O}((n_T + n_S)^3)$	$\mathcal{O}((n_T + n_S + d)^8)$	$\mathcal{O}((n_T + n_S)^3)$
収束レート	$\mathcal{O}_p(n_T^{-\frac{1}{2}} + n_S^{-\frac{1}{2}})$	$\mathcal{O}_p(n_T^{-\frac{1}{2}} + n_S^{-\frac{1}{2}})$	$\mathcal{O}_p(n_T^{-\frac{1}{2}} + n_S^{-\frac{1}{2}})$

n_S : 転移元ドメインでのデータの数

n_T : 転移先ドメインでの (ラベル無し) データの数

d : 入力データの次元

質的フィードバックを用いた比較バンディット

徐立元 (東大・理研AIP) 本多 淳也 (東大・理研AIP) 杉山 将 (理研AIP・東大)

目的 質的なフィードバックを扱うバンディット問題を提案

手法 フィードバックに基づく相対評価を行う比較バンディットとして定式化

結果 提案手法は比較バンディット手法の直接適用に比べてリグレットを低減

• 質的フィードバック

- 「累積報酬」が定義されない
- フィードバックの比較は可能



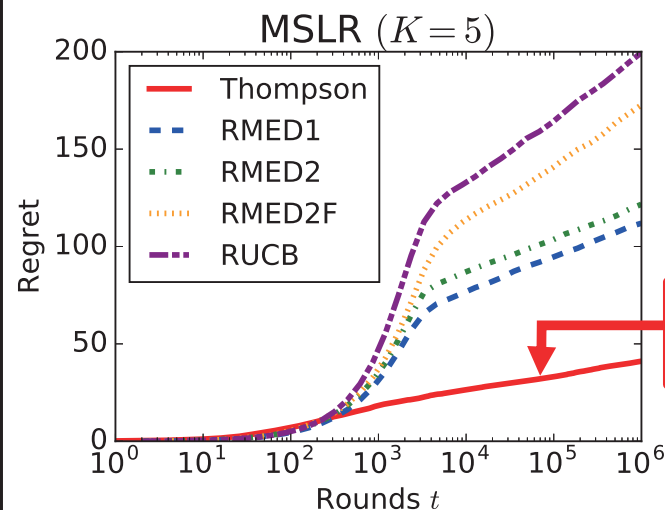
• 比較バンディット

- 腕間の比較を通じて最適腕を定義

フィードバック分布により比較を定義

• 実験結果

フィードバック情報を用いて、
既存手法より小さいリグレットを達成



提案手法

打ち切り行列の補完：天井効果への処方箋

手嶋 毅志(東大・理研) Miao Xu (理研) 佐藤 一誠(東大・理研) 杉山 将(理研・東大)

目的 行列の観測値が上限で打ち切られているとき、打ち切り前の行列を推定する

手法 真の行列は低ランクだと仮定し、成分間の依存関係を利用して補完

結果 打ち切り(多くの分野で現れる情報欠損)に対し、有効な行列補完法を提案

- (1)問題の可解性…核ノルム最小化による完全復元のための十分条件を提示
- (2)アルゴリズム…二乗ヒンジ損失による補完手法と、理論保証付き正則化を提案



天井効果を受けた成分も……

補完できる！

平均値シフトの性質の解析と課題

山崎 遼也 (京大) 田中 利幸 (京大)

目的

モード推定/クラスタリングのアルゴリズムである平均値シフト(MS)の性質を解明すること.

MSは適応的なステップサイズを用いた勾配上昇法でカーネル密度推定結果(KDE)の極大値を探索する.

手法

MSを特徴づけるMS関数の性質を詳細に調べ, y_{t+1} の近傍におけるKDE $\hat{p}(x)$ の挙動について解析する.

結果

- カーネル関数に関するある条件の下で,

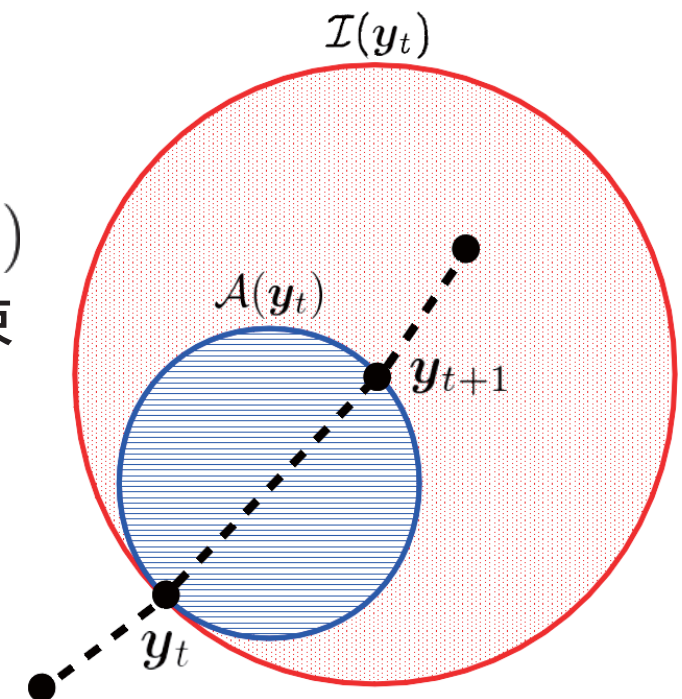
$$\hat{p}(x) \leq \hat{p}(y), \quad y \in \mathcal{I}(x)$$

$$\nabla \hat{p}(y) \cdot (y - x) \geq 0, \quad y \in \mathcal{A}(x)$$

- モード推定系列が「最も近い」モードに収束するための十分条件を導出.

課題

多次元で一般的なカーネル関数と適応的なステップサイズを用いたとき, モード推定系列が収束するか否か分かっていない.



敵対的生成ネットワークを用いた逆合成の経路探索

福島 真太郎 (株式会社トヨタIT開発センター) 本山 裕一 吉見 一慶 (東京大学物性研究所)

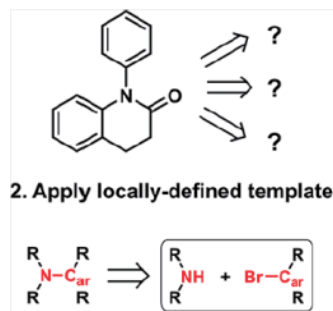
目的 有機分子の合成経路の外挿的な探索

手法 敵対的生成ネットワークによる分子の生成・反応式の増加

結果 既存手法に比べて探索の精度が向上

• 従来手法と問題点(Coley et al.(2017))

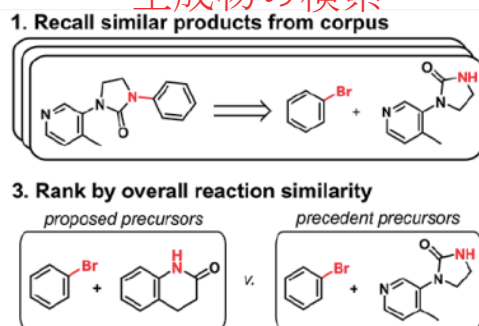
他手法に比べて高精度だが
内挿的な探索しかできない



2. 目的の生成物への
局所テンプレートの適用

C.W.Coley et al., ACS. Cent. Sci. 3 (12), 1237(2017)

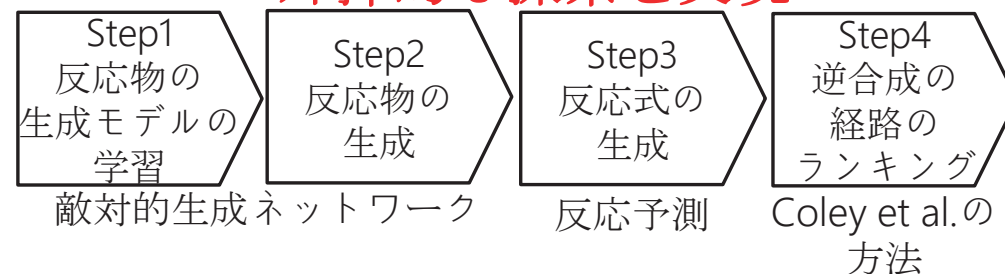
1. 目的の生成物と類似する
生成物の検索



3. 目的の生成物の
反応式ランキング

• 提案手法

敵対的生成ネットワークを用いて
外挿的な探索を実現



• 結果

モデル	トップ n に正解が含まれる割合					
	1	3	5	10	20	50
Liu et al.(2017) ベースライン	35.4	52.3	59.1	65.1	68.6	69.5
Liu et al.(2017) Seq2Seqモデル	37.4	52.4	57.0	61.7	65.9	70.7
Coley et al.(2017) クラス分割なし	37.3	54.7	63.3	74.1	82.0	85.3
提案手法	39.4	58.1	67.4	76.3	84.3	86.0

新聞印刷工場のスマートファクトリー化

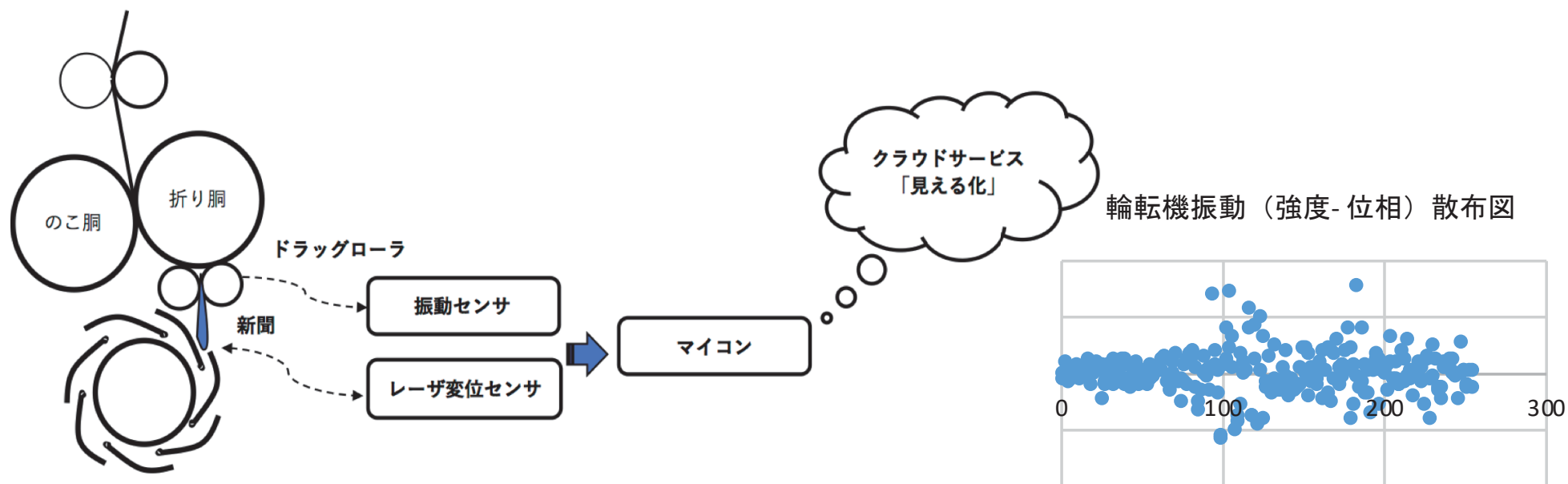
六車 浩二（九産大） 田中 康一郎（九産大）

目的 AI/IOTを利用したスマートファクトリー化追及

手法 「Low Cost」 「Small Start」 「Easy Maintenance」 アプローチ

結果 現場の知見に基づくアイデアを実現

- 新聞印刷輪転機で折畳部の新聞通過タイミングに着目した不良紙発生予測（振動値散布図学習モデルと散布図カメラ画像から危険域リアルタイム予測）



広告配信における選択バイアスの学習とその補正

鈴木 真矢, 安部 斉志, 平澤 謙章, 盛岡 玲資
ヤフー株式会社

目的

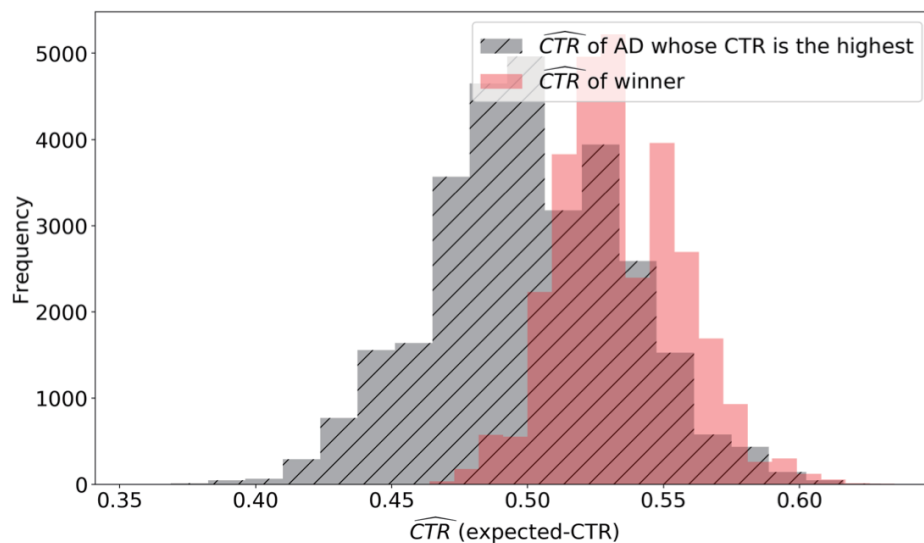
ランキングによって生じる広告クリック確率のバイアスを補正する

手法

バイアスの説明変数を用いて機械学習モデルを構築

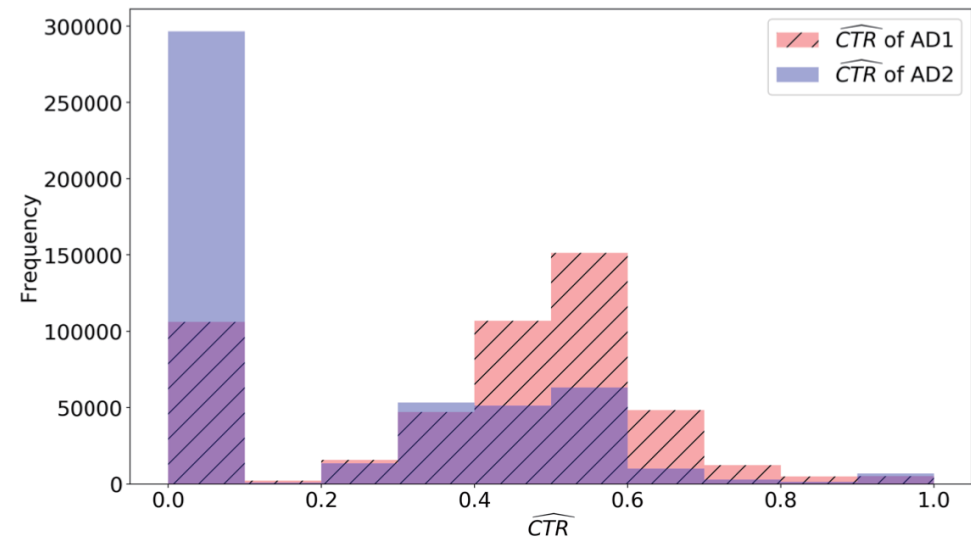
結果

クリック確率の予測精度が向上し, またバイアスの減少が確認された



Ranking Bias

予測が過大評価された広告が選ばれやすい



Negative Bias

予測が過小評価されると配信機会を失い下振れ続ける

メタゲノム中の プラスミド配列検出システムの開発と検証

原田 誠史 鈴木 治夫 富田 勝 (慶應義塾大学)

目的

手法

結果

高次元におけるk近傍法の漸近的性質とバイアス補正

中山 優吾 (筑波大)

目的 高次元における次元の呪いを受けないk近傍法の提案

手法 ノイズ掃き出し法による固有ベクトルを用いた次元削減とバイアス補正

結果 次元削減後は従来のk近傍法の性能を理論的・数値的に向上させられる

- 高次元小標本 ($d > n$) データが様々な現場で観測される.

→ マイクロアレイデータ・画像データ etc.

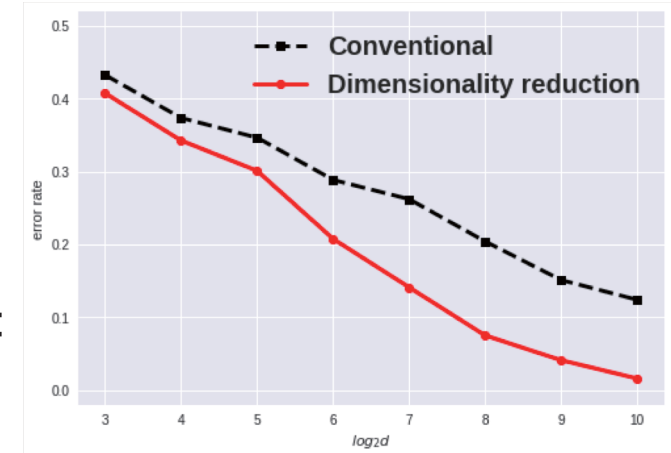
- k近傍法は次元の呪いの影響を受けて精度が悪化.

- 主成分分析によって次元削減が有効 (図を参照).

→ **標本固有ベクトル**は高次元で**不一致性**を持つ:

$$\mathbf{h}_i^T \hat{\mathbf{h}}_i = (1 + \delta_i)^{-1/2} + o_P(1) \quad d, n \rightarrow \infty$$

$\Sigma = \Sigma_{i=1}^d \lambda_i \mathbf{h}_i \mathbf{h}_i^T$ 共分散行列, $S = \Sigma_{i=1}^{n-1} \hat{\lambda}_i \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^T$, 標本共分散行列, $\delta_i = \lambda_i^{-1} \Sigma_{j=m-1}^d \lambda_j / (n-1)$, m はある定数



- Yata and Aoshima (2012, JMA) が提案した **“ノイズ掃き出し法”** を固有ベクトル推定に応用することで精度を向上させる.

順序付けICAによる実データ分析

松田 源立 (東京大学)

目的 源信号を一意に順序付けして推定可能な独立成分分析(ICA)の開発と応用

手法 順序付け可能な目的関数の導出および並列化による大域的最適化

結果 ICAによるEEG信号解析の性能向上

- 独立成分分析(ICA)
 - 教師無しの源信号線形分離手法の一種
 - 源信号が非ガウス性を持ちかつ互いに独立に発生していると仮定
 - 音源分離、画像分離、特徴抽出、EEG信号解析等の様々な分野で応用
- 順序付けICA
 - 一意に順序付けして源信号を推定するICA手法
 - 並列化による大域的最適性の向上
 - 情報量基準の利用によりガウスノイズの除去が容易
- 実データ(EEG信号)でのシミュレーションにより順序付けICAの有効性を実証

深層ガウス過程により表されたダイナミクスの安定性解析

福永 修一 (東京都立産業技術高等専門学校) 佐藤 訓志 (大阪大学)

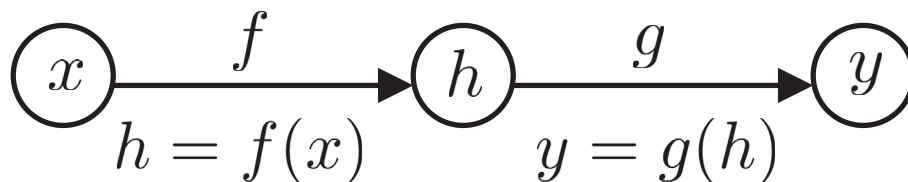
目的 深層ガウス過程により表された離散時間ダイナミクスの安定性解析

手法 各階層毎の解析から全階層の伝搬を考慮した出力の時間発展の上界を計算

結果 深層ガウス過程により表された離散時間ダイナミクスの安定化領域を明示的に与えることが可能

キーワード 深層ガウス過程 $\Rightarrow$ ガウス過程に**深層構造**を導入

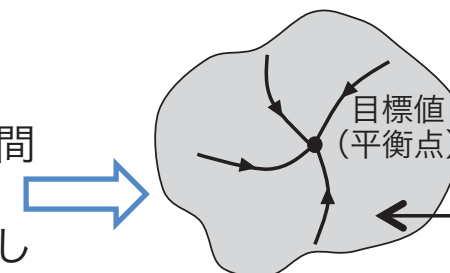
例：2層の深層ガウス過程



入力 x から出力 y における
 x に関する安定化領域

提案アルゴリズム

1. 各階層毎に出力の時間発展に関する上界をそれぞれ計算
2. これらを統合することで全階層の伝播を考慮した入出力間の時間発展に関する上界を獲得
3. 出力の時間発展が目標値（平衡点）に対して単調に収束していくための上界条件が得られる
4. この条件を満たす入力の初期値の集合を求める



最終的にシステムの
安定化領域が求まる

安定化領域：
目標値（平衡点）に漸近収束できる初期値の集合

グラフ埋め込みの次数補正とその応用

田中 卓磨 (京大・理研AIP) 奥野 彰文 (京大・理研AIP) 下平 英寿 (京大・理研AIP)

目的

次数の高いノードが他のノードとリンクを張りやすくなるようグラフ埋め込みの既存モデルを修正しリンク予測精度などの向上を目指す。

手法

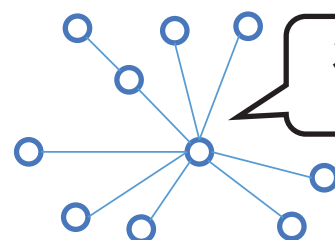
既存のグラフ埋め込みの確率モデルに次数補正項を加える。

結果

実験での比較により、提案モデルの性能が既存モデルを上回ることを示した。

グラフ埋め込みの確率モデル $w_{ij} \mid \mathbf{x}_i, \mathbf{x}_j \sim \text{Po}(\mu(\mathbf{x}_i, \mathbf{x}_j))$

■ 既存モデル $\log \mu(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{f}_{\text{NN}}(\mathbf{x}_i), \mathbf{f}_{\text{NN}}(\mathbf{x}_j) \rangle + \log \alpha$



次数が高いノードは他のノードと接続しやすい
(次数が低いノードは他のノードと接続しにくい)

図: グラフの特定のノードの次数は大きい

■ 提案モデル $\log \mu(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{f}_{\text{NN}}(\mathbf{x}_i), \mathbf{f}_{\text{NN}}(\mathbf{x}_j) \rangle + \log \beta_i + \log \beta_j$

次数補正項

(※同様の補正項(バイアス項)は推薦システムなどでも有効性が知られている)

次数補正項 β_i の推定には、(1) パラメータとみなし尤度をminibatch SGDで最大化、(2) 観測値の次数と比例させる、(3) NNを用いて学習する、の3通りを用いた。

レビュー者特性を考慮した多次元項目反応モデルの提案

八木 嵩大（電気通信大学） 宇都 雅輝（電気通信大学）

目的

レイティングデータからレビュー者特性（甘さ/厳しさなど）を考慮して多次元スコアリングを行う潜在変数モデル（項目反応モデル）の提案

手法

レビュー者特性パラメータを付与した項目反応モデルのパラメータを多次元に拡張し、MCMCによるパラメータ推定法を提案

結果

- ・ シミュレーション実験によりモデルとパラメータ推定の妥当性を確認
- ・ 実データ実験により提案モデルの妥当性を確認

提案モデル

評価対象 i に対しレビュー者 r が評価観点 j について評点 k を与える確率は次式

$$P_{ijrk} = P_{ijrk-1}^* - P_{ijrk}^* \quad \begin{array}{l} \text{スコアの次元数} \\ \text{評価対象 } i \text{ の } l \text{ 次元目のスコア} \end{array}$$

$$\begin{cases} P_{ijrk}^* = \left\{ 1 + \exp \left(-\alpha_r \left(\sum_{l=1}^L \alpha_{jl} \theta_{il} - \beta_{jk} - \varepsilon_r \right) \right) \right\}^{-1} & k = 1, \dots, K-1 \\ P_{ijr0}^* = 1 & \text{レビュー者 } r \text{ の一貫性} \\ P_{ijrK}^* = 0 & \text{レビュー者 } r \text{ の厳しさ} \end{cases}$$

$$\begin{array}{l} \text{評価観点 } j \text{ の } l \text{ 次元目の識別力} \\ \text{評価観点 } j \text{ において評点 } k \text{ を得る困難度} \end{array}$$

半解析的ブートストラップ法とその応用

小渕智之（東工大情報理工） 樺島祥介（東工大情報理工）

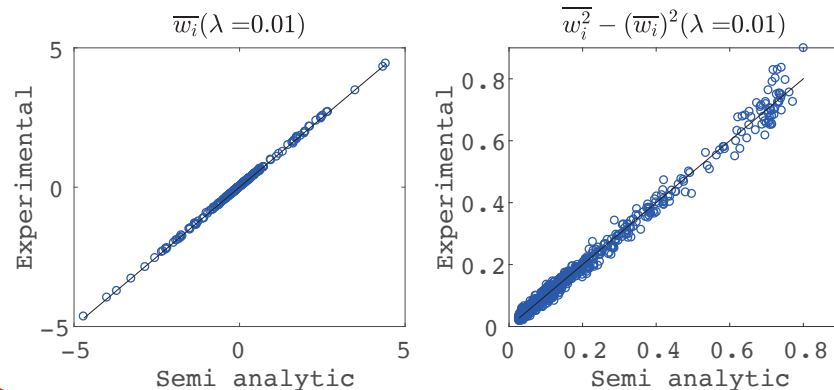
目的 複雑な推定・確率モデルにおける推定量に信頼度（信頼区間）をつけたい

手法 ブートストラップ法の枠組みをつかう．その計算量を統計力学的近似で削減することで、複雑なモデルにも使えるようにする．

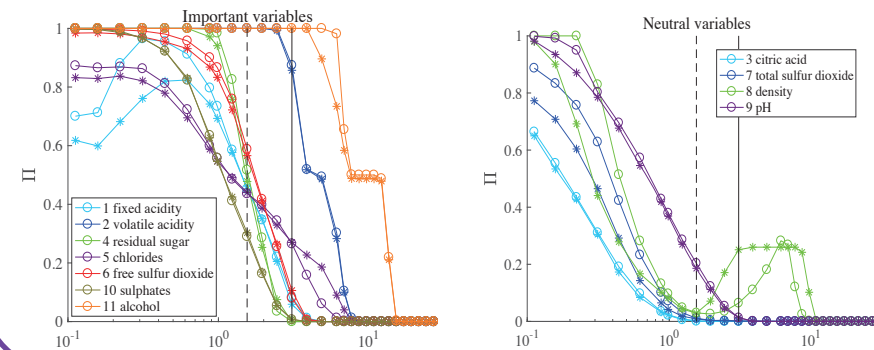
結果 LASSOにおいて低計算量で精度の高い近似公式を提案．
Stability selection(SS)との併用で、現実的な変数選択法を構成．

- データをブートストラップ法でシャッフル： $D \Rightarrow D_c$
- LASSO推定量 $\hat{w}(\lambda, D_c)$ の c に関する分布を計算： $P(\hat{w})$

人工データでの比較



ワインクオリティデータにおけるSS



深層ベイズ学習に基づく単チャネル音源分離の教師なし学習

坂東 宜昭 佐々木 洋子 (産総研)

目的

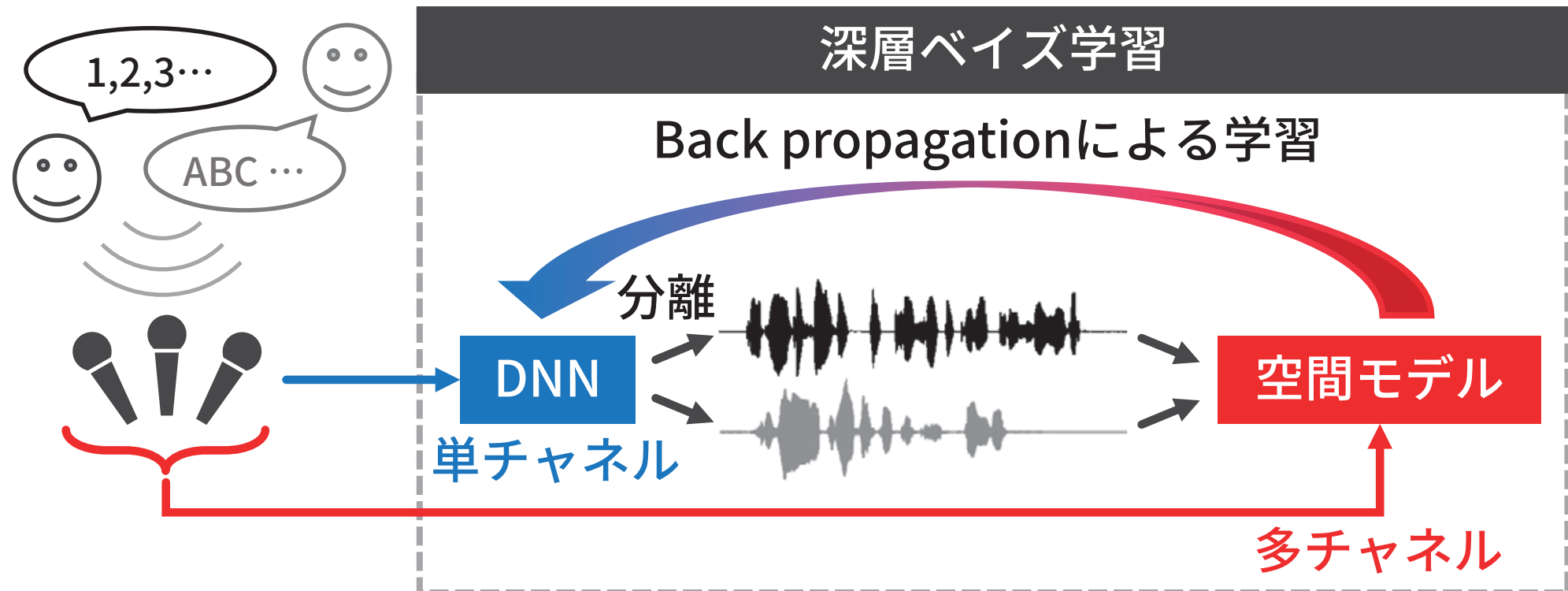
単チャネル音源分離を(正解信号を使わずに)混合音のみで学習したい

手法

多チャネル信号を学習データに，空間モデルに基づくコスト関数で学習

結果

方向に偏りがある学習データからでも，頑健に音源分離を獲得



ボルツマンマシンによる強化学習

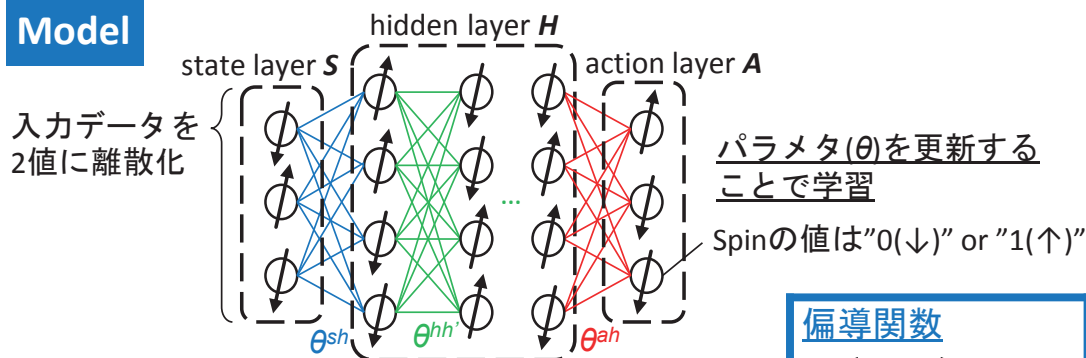
内垣内 洋 (日立製作所) 山岡 雅直 (日立製作所)

目的 Deep Boltzmann machine(DBM)を用いた学習の汎用性及び **学習速度**を評価。

手法 入力値の離散化 + DBMの熱平衡における物理量を評価関数として利用。

結果 3種の問題を求解可能. パラメタの **更新回数**はNeural network(NN)の1/3~2/3.

Model



行動価値関数

$Q(s, a; \theta)$

$$= \sum_{\substack{s_i \in S \\ h_j \in H}} \theta_{ij}^{sh} s_i \langle h_j \rangle + \sum_{\substack{a_i \in A \\ h_j \in H}} \theta_{ij}^{ah} a_i \langle h_j \rangle + \sum_{\{h_i, h_j\} \subseteq H} \theta_{ij}^{hh'} \langle h_i h_j \rangle - \frac{1}{\beta} \sum_h \mathbb{P}(h|s, a; \theta) \log \mathbb{P}(h|s, a; \theta)$$

偏導関数

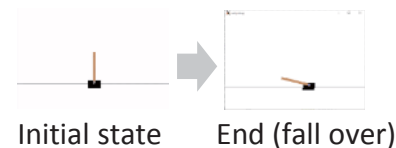
$$\frac{\partial Q(s, a; \theta)}{\partial \theta_{ij}^{sh}} = s_i \langle h_j \rangle$$

$$\frac{\partial Q(s, a; \theta)}{\partial \theta_{ij}^{ah}} = a_i \langle h_j \rangle$$

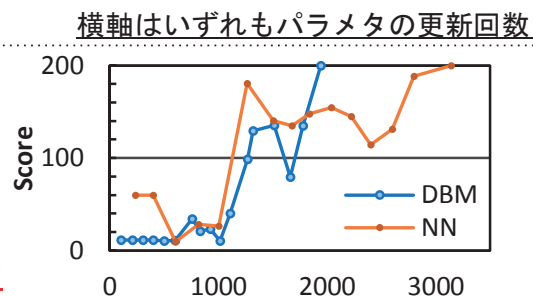
$$\frac{\partial Q(s, a; \theta)}{\partial \theta_{ij}^{hh'}} = \langle h_i h_j \rangle$$

今後 更新回数の削減手法の検討, 他の適する問題の探索

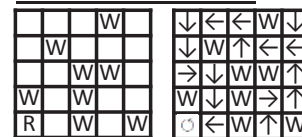
Cartpole



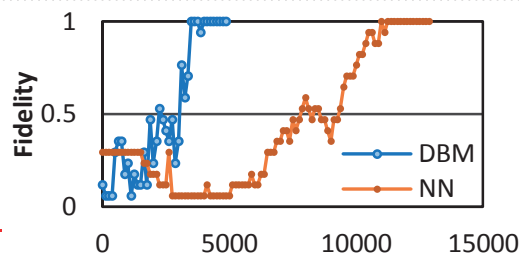
結果 約2/3の更新回数で学習



Maze traversal



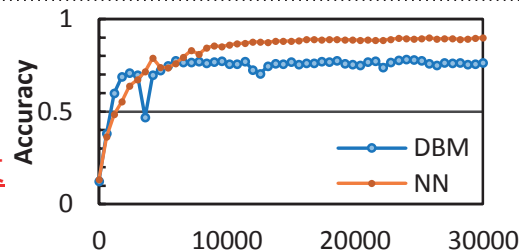
結果 約1/3の更新回数で学習



MNIST

- 5x5に縮小
- 教師付き学習

結果 約1/3の更新回数で収束
Accuracy: ~13%



転移学習における実験計画的ドメイン選択

松井 孝太（理研AIP） 熊谷 亘（理研AIP）

福地 一斗（理研AIP） 竹内一郎（名工大・NIMS・理研AIP）

目的 目標ドメインのタスク性能に寄与する元ドメインを**低コスト**で選びたい

手法 元ドメインに対する重要度の適応的な決定問題として定式化

結果 **目標ドメインの期待リスク上界に基づいた能動学習**アルゴリズムを提案

目標ドメインの期待リスク上界（サンプル近似）

$$R_T(h) \leq \underbrace{\sum_{k=1}^K \alpha_k \hat{R}_{S_k}(h)}_{\text{K個の元ドメインの経験リスクの凸結合}} + \sum_{k=1}^K \alpha_k \left(\underbrace{(\text{Disc}(\hat{P}_T, \hat{P}_{S_k}) + \mathcal{O}\left(\frac{1}{\sqrt{n_T}} + \frac{1}{\sqrt{n_{S_k}}}\right))}_{\text{元ドメインと目標ドメインの経験分布の不一致度}} + \underbrace{\text{Diff}(f_T, f_{S_k})}_{\text{元ドメインと目標ドメインのラベル関数の不一致度}} \right)$$

$$= F(\alpha)$$

$F(\alpha)$ に基づいてドメインの重要度 α を適応的に決定する能動学習を行う

複数データセットを用いたGANによる データ拡張手法の検討

山口 真弥 村松 沙那恵 江田 毅晴
(NTT Software Innovation Center)

目的

少量の学習データで高精度な深層学習

手法

一つのGANを複数のデータセットで同時に学習

結果

1/5の学習データ量で全データ量と同等の精度を達成 (Re-IDタスク)

・ 深層学習は学習データ整備がツライ！！！！！！！！！！！！！！！！！！！！

・ 実用化案件ではデータの収集・アノテーションがボトルネック

・ ここをなんとかしないと何も進まない

・ GANによるデータ拡張

・ 最近結構取り組まれている

・ でも, GANを高精度に学習するためにそもそもデータがたくさん必要

・ 提案: GANのマルチデータセット学習

・ 所望のデータが足りないなら, あるものを使えばいいじゃない

・ 通常のcGANよりも高品質(FID)かつ学習に役立つデータ生成を実現

表. Market1501 (751(学習)クラス) における
クラスごとの学習データ数に対する平均目的モデル精度

# Samples per class	$n = 4$	$n = 20$
Base Line	0.2825	0.5045
cGAN	0.4189	0.6162
Market-CIFAR10	0.5586	0.6208
Market-CUHK03	0.4589	0.5980

マルチエージェント自動交渉における メディエーションとベイズ統計の関係について

山崎 啓介（産総研） 藤田 桂英（農工大）

目的

複雑な交渉プロセスにおける合意案の最適性を評価

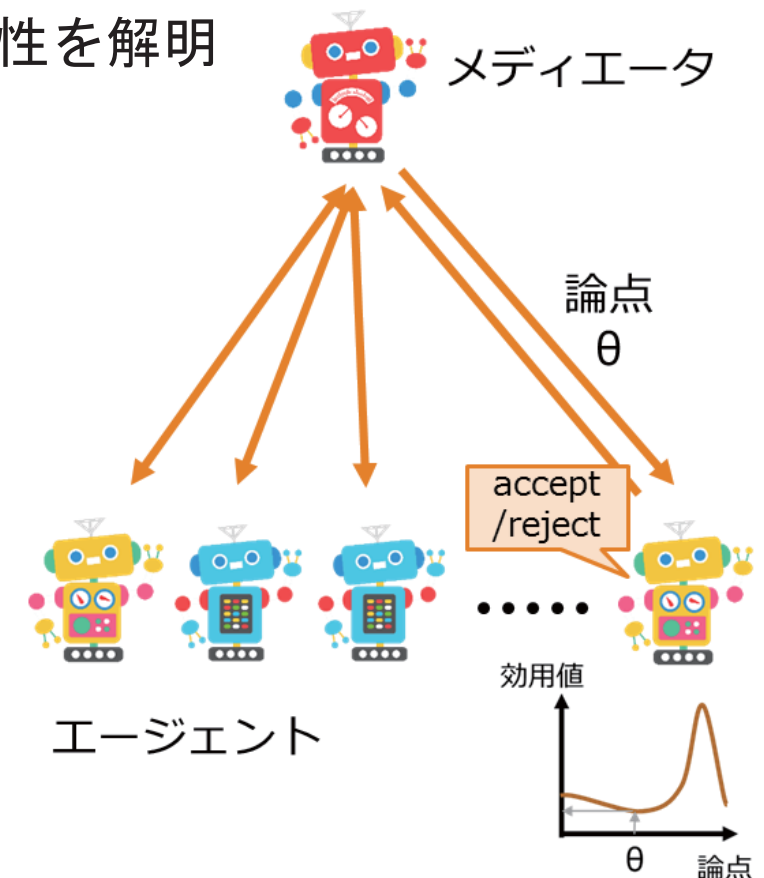
手法

交渉プロセスをMCMCサンプリングとして定式化

結果

事後分布の漸近解析により合意案の収束特性を説明

- 自動交渉：2人以上のエージェントが個々の利益に基づき交渉を行い合意を形成する
- 提示される条件（論点）に対し、効用関数で定められる選好（効用値）に従い、合意するか拒否するかを決定する（戦略）
- マルコフ連鎖モンテカルロ法を用いた合意案探索を行うプロトコルとエージェントモデルを定式化し、ベイズ統計との関連性を示す



Adaptive Intention Learning

比嘉 亮太 (NEC)

目的 環境変化を考慮した意図学習の考案

手法 報酬(意図)設計に環境変化の繰り込む手法

結果 報酬設計により環境変化に追従可能になった

- 環境変化の概要をこのページ、概要の図をこのページへ
- 報酬設計、未知のダイナミクス学習(別ポスター)、メモリー活用など

深層学習におけるskip結合とその表現集合

○長瀬 准平(芝浦工業大学大学院理工学研究科)

目的 skip結合の比較、ResNetとDenseNetの比較。

手法 Affine関数の合成に注目し、「写像の圧縮」を用いてNNの関数集合を調べた。

結果 パラメータの制限の一つとしてResNetとDenseNetの新たな対応を得た。

- Affine関数の合成の性質に注目し、「写像の圧縮」と「正則圧縮同値」を定義した。
- 上記二つの概念を用いることで、ResNet[He+ 2015]とDenseNet[Huang+ 2016]の対応関係を新たに説明する、パラメータの制限による解釈を数学的に導出した。
- 「写像の圧縮」は制限写像の自然な拡張であり、純粋数学との関連が期待される。
- 「正則圧縮同値 $\approx_{\text{com}}$ 」は、関数の集合に対して合成演算から導かれる同値関係であり、関数解析などの分野での発展が期待される。

$$\mathcal{A}^{d_{L+1} \times \sum_{i=0}^L d_i} \circ \left(f \circ \mathcal{A}^{d_L \times \sum_{i=0}^{L-1} d_i}, \text{id} \right)^T \circ \dots \circ \left(f \circ \mathcal{A}^{d_1 \times d_0}, \text{id} \right)^T$$

$$\approx_{\text{com}} \quad (\mathcal{A}^{d_{L+1} \times d_L} \circ f + \mathcal{A}^{d_{L+1} \times d_L}) \circ \dots \circ (\mathcal{A}^{d_2 \times d_1} \circ f + \mathcal{A}^{d_2 \times d_1}) \circ \mathcal{A}^{d_1 \times d_0}$$

バイオ配列のための編集不変ニューラルネットワーク

小出 智士, 河野 圭祐, 沓名 拓郎 (豊田中央研究所)

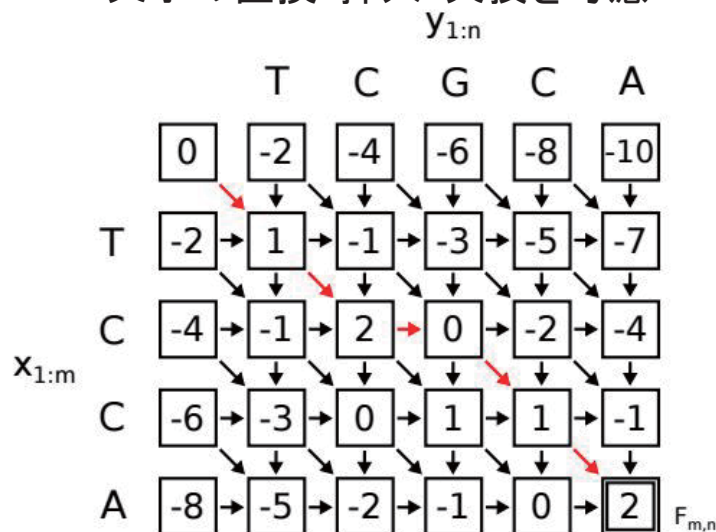
目的 バイオ配列の進化の特徴を捉えた「編集不変性」をもつ NN の提案

手法 微分可能な配列アライメントを CNN の内積の代替として利用 (CNN の一般化)

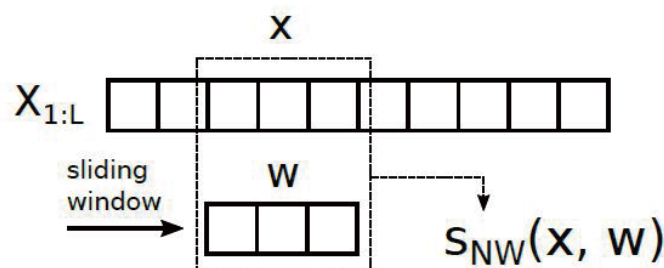
結果 少数のデータのケースにおいて CNN よりも高精度を実現

配列アライメント
(Needleman-Wunch)

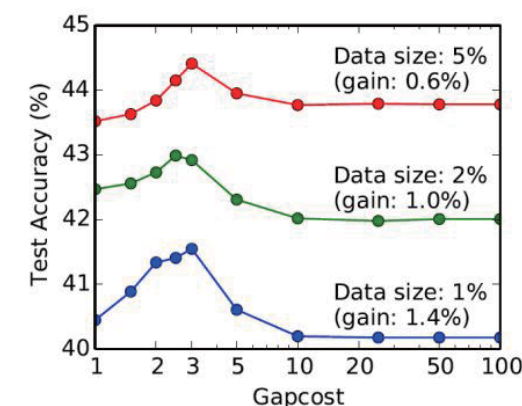
→ 文字の置換・挿入・欠損を考慮



配列アライメントで
CNN の内積を代
替



少数データほど有利



その他の結果: CNN と star-free regular expression の関係についても考察
詳細: Koide et al., Neural Edit Operations for Biological Sequences, NIPS'18 (to appear)

動径分布関数のカーネル密度推定

徳田 悟 (産総研) 平田 秋彦 (早大・産総研)

目的

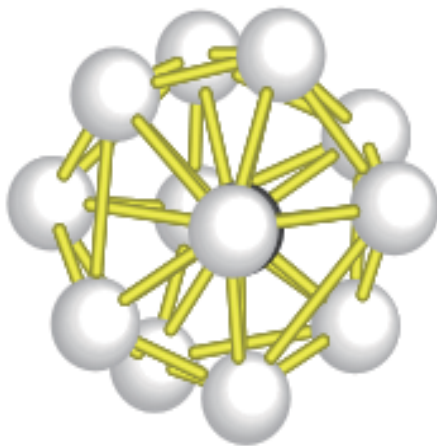
アモルファス金属の局所構造の動径分布関数を導出

手法

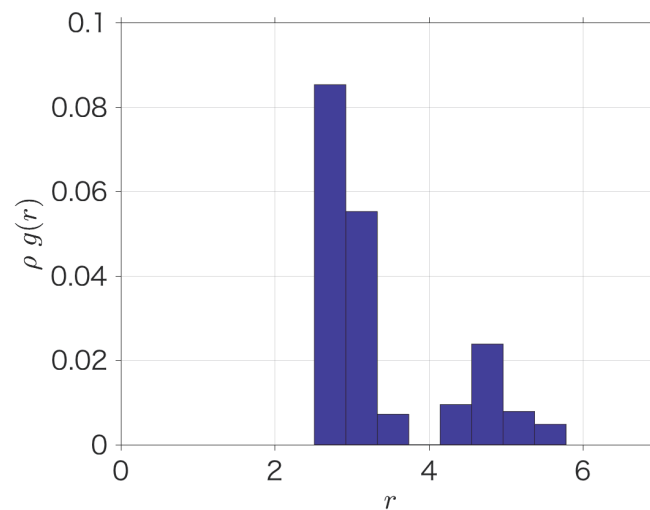
熱振動の調和近似、隠れ変数の周辺化

結果

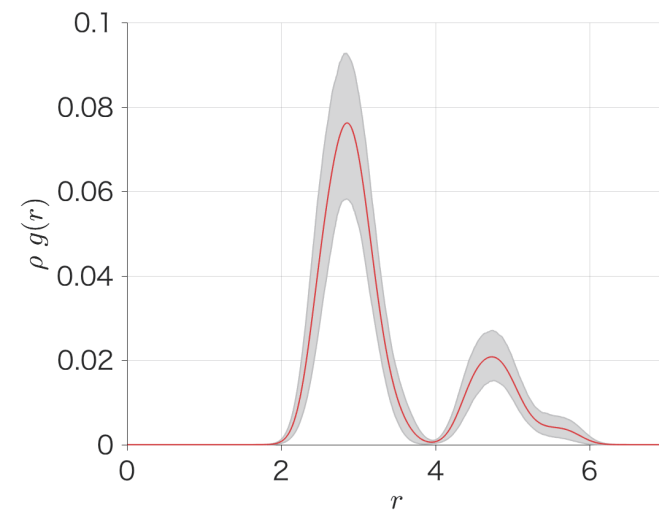
動径分布関数をカーネル密度推定の表式として取得



対象データ
(二十面体クラスタ)



従来法
(ヒストグラム)



提案法
(カーネル密度推定)

カーネル平均埋め込みによる尤度なし関数を用いた 共変量シフト下での回帰手法

木佐森 慶一 (産総研 / NEC), 山崎 啓介 (産総研)

目的

解析的に書けない関数を用いた回帰と、その共変量シフト対応

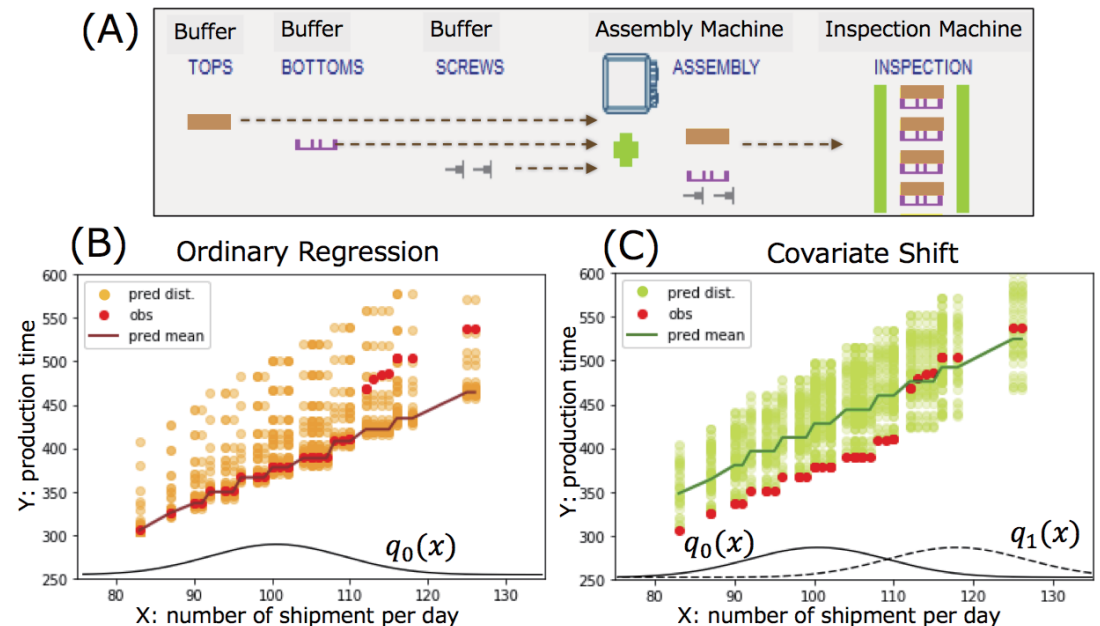
手法

重要度重み付きカーネルを用いたカーネル平均埋め込み

結果

離散的なシミュレータを回帰関数として共変量シフトした回帰に成功

- 応用例: 生産工程を模擬するシミュレータを回帰関数とした製品生産効率予測
 - 解析的に書けない関数 $y = F(x; \theta)$
 - x : 1日の製品生産台数
 - y : 製品を x 個作るための時間
 - θ : 各工程の作業時間
- 期待される効果:
 - 外挿的な領域での高精度な予測
 - シミュレーションユーザにとって、
解釈可能なパラメタの提示



強低ノイズ条件下識別問題に対する 確率的勾配降下法の線形収束性

二反田 篤史 (東大) 鈴木 大慈 (東大・理研AIP)

目的 強低ノイズ条件下識別問題での確率的勾配降下法 (SGD) の収束性を研究

手法 SGDの収束・安定性理論を援用し損失ではなく識別誤差を直接解析

結果 SGDにより期待識別誤差が線形収束 (損失関数値の収束率は劣線形)

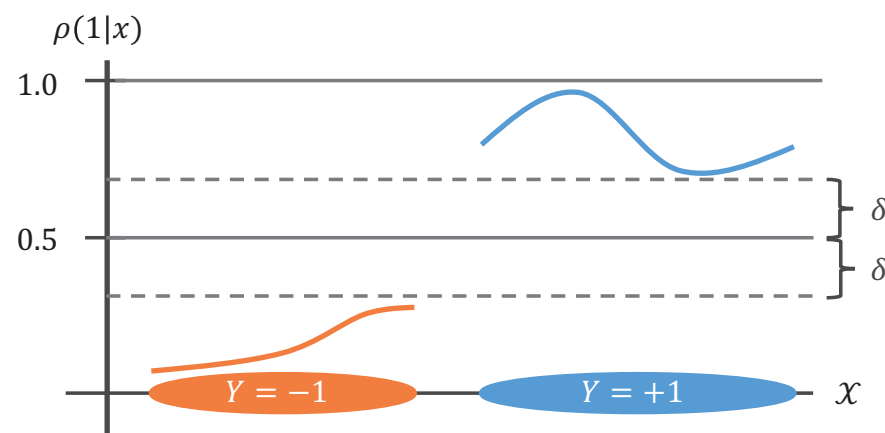
- 正則化項付き二値識別問題 : $\mathcal{X}$: 特徴空間, $\mathcal{Y} = \{-1, 1\}$, $\mathcal{H}_k$: RKHS (仮説空間)

$$\min_{g \in \mathcal{H}_k} \left\{ \mathcal{L}_\lambda(g) = \mathbb{E}_{(X,Y)} [l(g(X), Y)] + \frac{\lambda}{2} \|g\|_{\mathcal{H}_k}^2 \right\}$$

- 強低ノイズ条件** : $|\rho(1|x) - 0.5| > \delta$, ($\rho(1|x)$: ラベル1の条件付確率)

- SGDの**期待識別誤差**の収束率 : ϵ : 求解誤差

$$O\left(\frac{1}{\delta^2 \lambda^2} \log\left(\frac{1}{\epsilon}\right)\right) \quad \text{損失の収束より圧倒的に高速な線形収束性}$$



類似探索に基づく音源分離

澤田 宏 (NTT) 青山 一生 (NTT)

目的

2秒程度以下の短い観測信号に対して高精度に音源分離を行う。

手法

クリーン信号DBを事前に構築. 実行時には類似探索結果でパラメータ更新.

結果

Close条件では高精度に分離. Open条件では従来技術並み. DB構築課題.

最小化すべき目的関数

$$\sum_{t=1}^T \sum_{n=1}^N \sum_{f=1}^F \left(\frac{|y_{f t n}|^2}{v_{f t n}} + \log v_{f t n} \right) - 2T \sum_{f=1}^F \log |\det \mathbf{W}_f|$$

分散パラメータ $\{v_{f t n}\}$ の推定がポイント

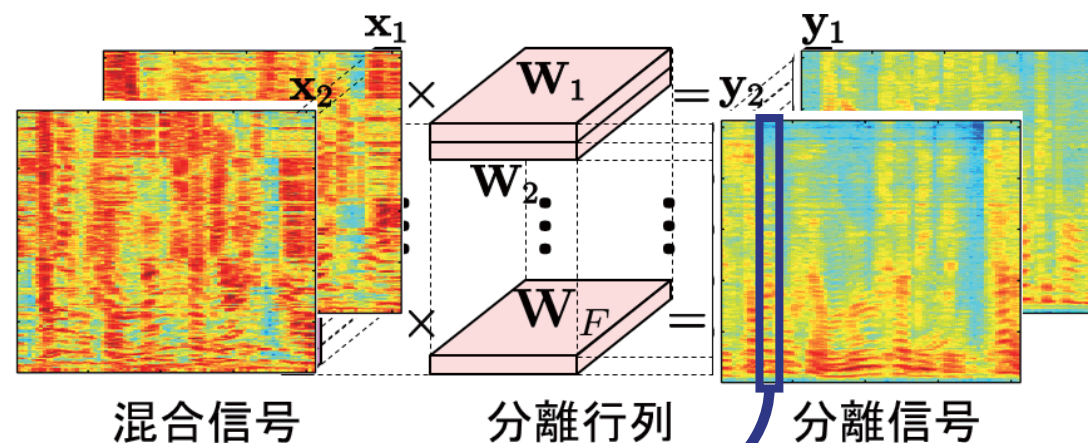
f 周波数

t 時間

n 信号id

従来技術

制約なし(ICA)
時間同期(IVA)
低ランク構造(ILRMA)
教師あり学習(DNN)



提案技術

類似探索

クリーン信号
データベース(DB)



識別と再構成を行うマルチタスク Deep Convolutional Neural Network の検討

小林 源太 (電通大) 庄野 逸 (電通大)

目的 再構成と識別課題を行うDCNNの相互課題の関連性を検討する

手法 2種類のDCNNに再構成損失を制約として加えて、学習の推移を観察する

結果 スキップコネクションを用いたDCNNは過学習の度合いが比較的小さかった

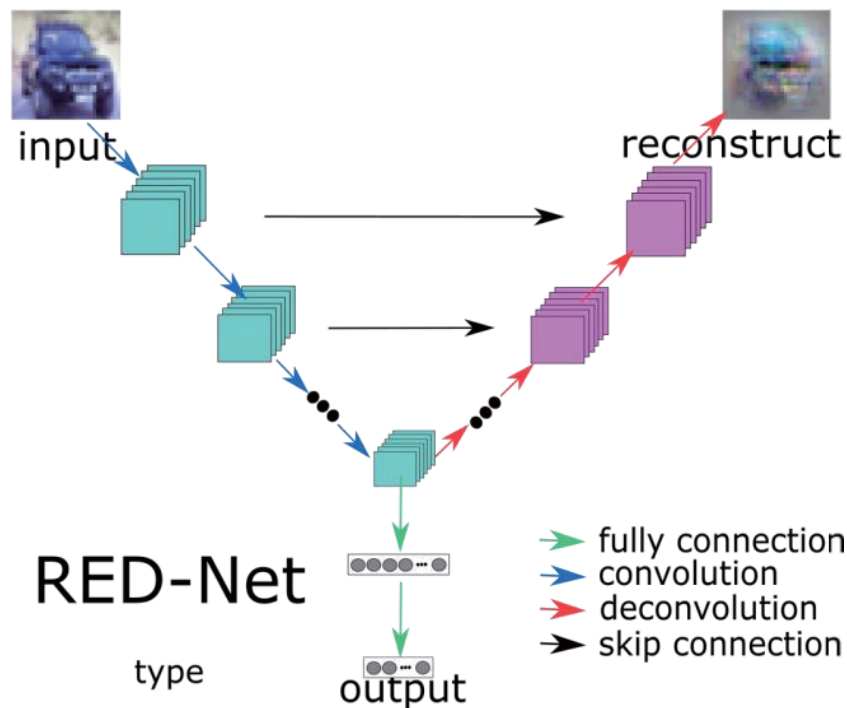


図1 スキップコネクションを用いたDCNN

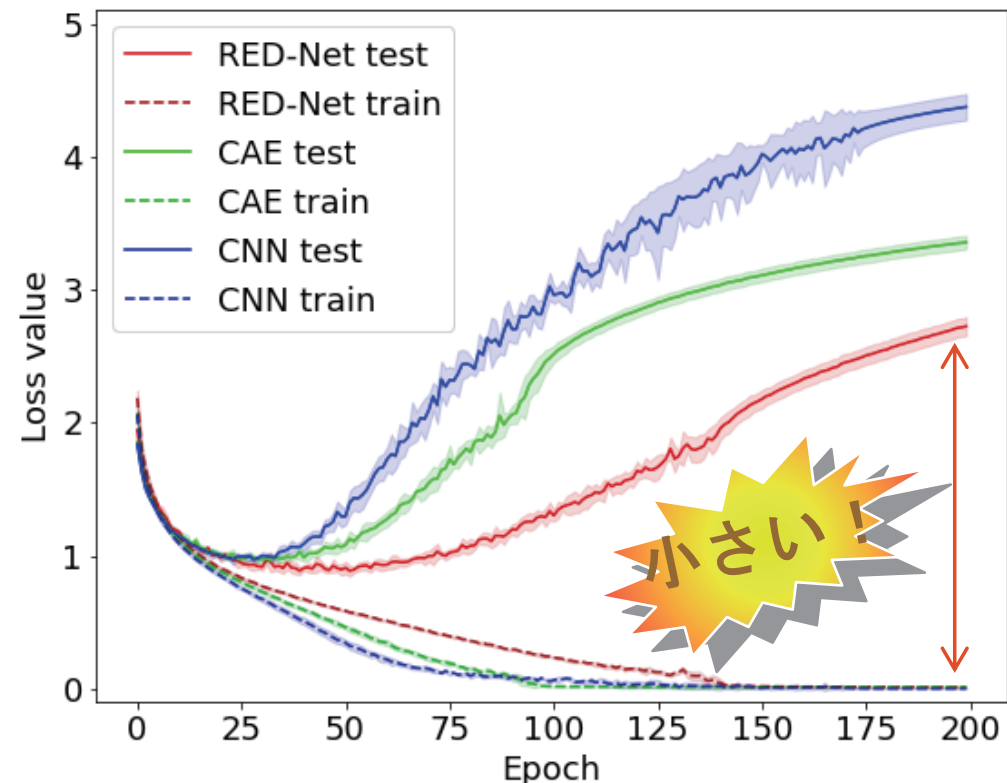


図2 各モデルの損失の推移

マルチモーダルな特徴量を用いた TVCM 動画の評価予測

熊谷 雄介 田原 將志 黄 橙白 道本 龍 (株式会社博報堂)

目的

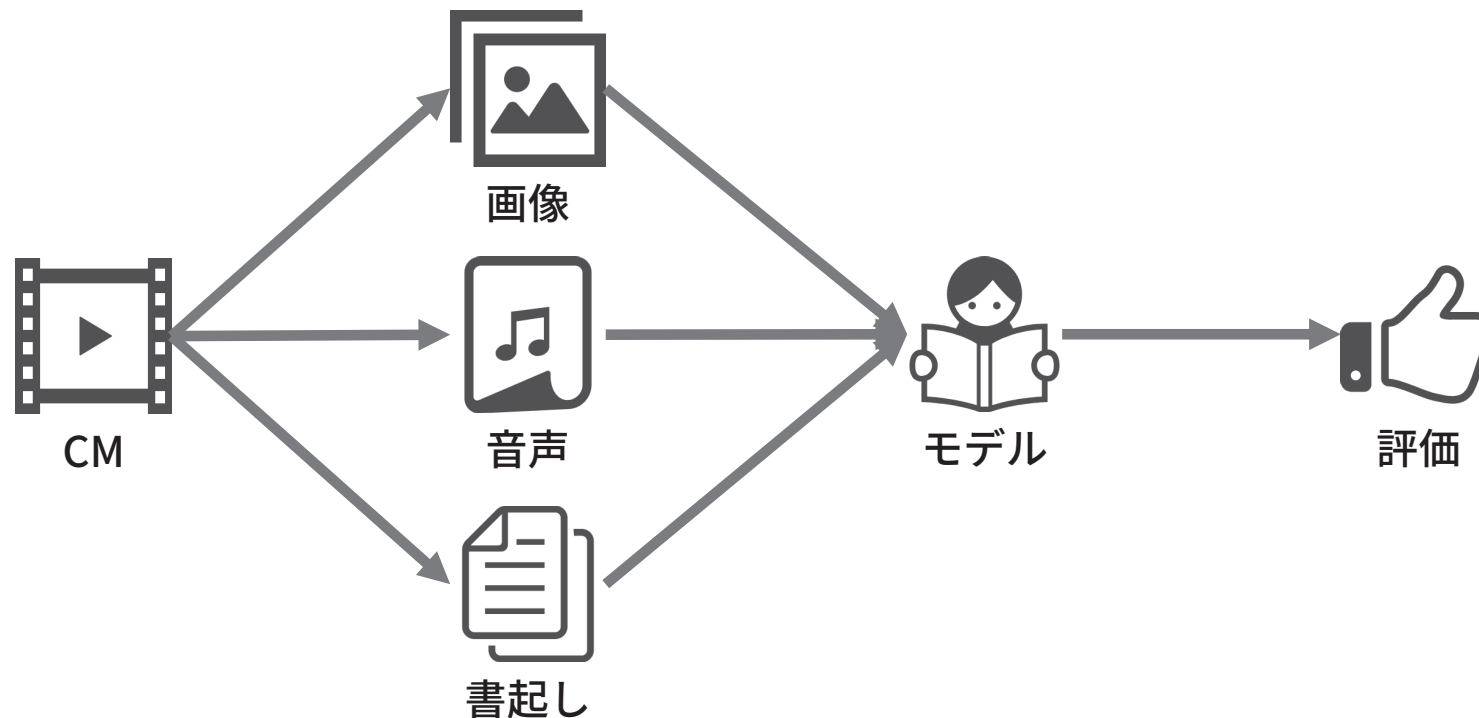
生活者による TVCM に対する評価を動画素材から推定したい

手法

マルチモーダルな特徴量を抽出して学習・予測

結果

R^2 0.63 - 0.70 で完全に未知の企業の TVCM 評価予測を実現



構造が時間に依存して変化するデータの埋め込み

布施拓馬(早大) 日野英逸(統計数理研) 赤穂昭太郎(産総研) 村田昇(早大)

目的

構造が時間に応じて変化するデータの連続性を考慮した座標を獲得する

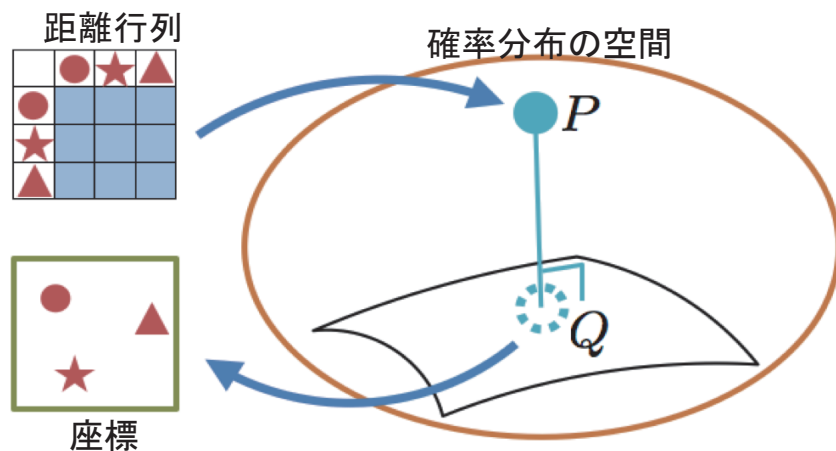
手法

埋め込み座標を分布の空間でスムージングする

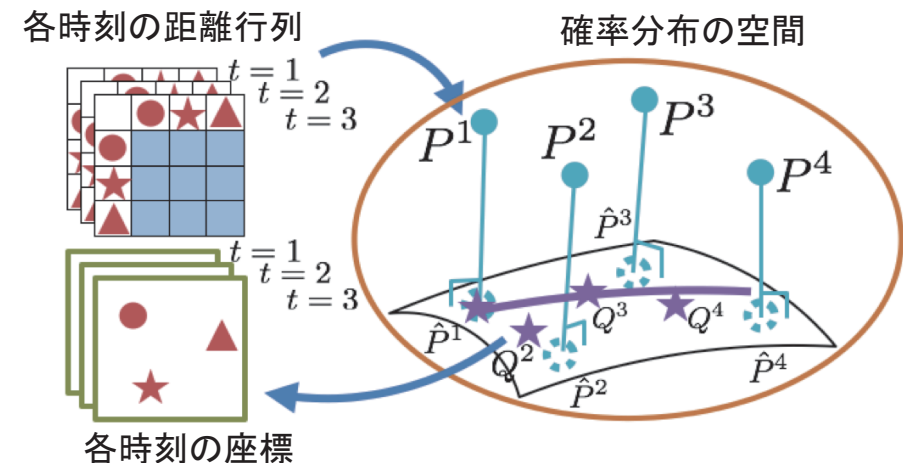
結果

提案手法が計算量の少ない手法に帰着することを証明した

一般的なSNE : $L = KL(P\|Q) = \sum_i KL(P_i\|Q_i)$



提案手法 : $L = wKL(P^t\|Q^t) + (1-w)KL(Q^{t-1}\|Q^t)$



- 1つのオブジェクトに対し, 2種類の座標を獲得する
→近傍確率も2種の座標の内積で定義する(word2vecと同じ形式)
- 損失関数に一時刻前の情報を導入する(分布の空間でのスムージング)
- 情報幾何的な解釈より, 提案手法は簡単なアルゴリズムに帰着する

Weighted Jensen-Shannon divergence 規準の ランダムフォレストを用いた条件付き分布の推定

五十嵐 里紗(早大) 關 翼人(早大)

日野 英逸(統計数理研) 赤穂 昭太郎(産総研) 村田 昇(早大)

目的

回帰分析における条件付き分布のノンパラメトリック推定

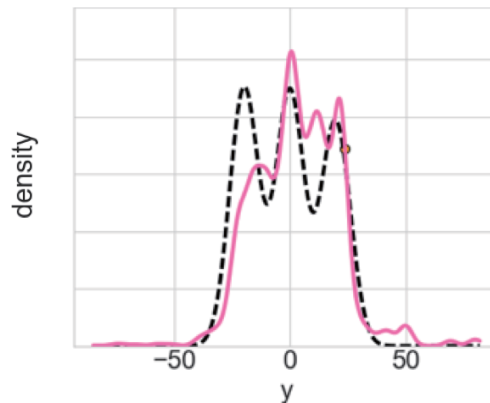
手法

Weighted Jensen-Shannon divergence を分割規準にしたランダムフォレスト

結果

人工データ実験において推定精度を定量的に評価した

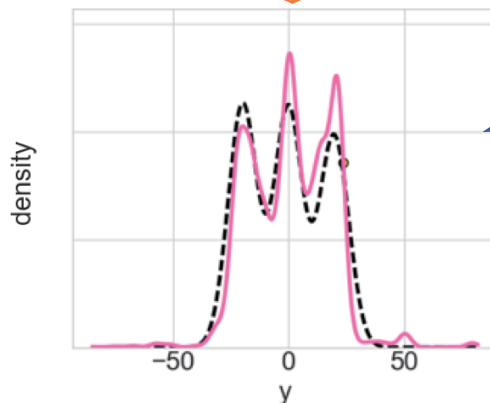
既存手法^[1]



回帰分析

条件付き分布 $p(y|x)$ を
ノンパラメトリックに推定したい

提案手法



推定に適した

ランダムフォレストの分割規準
として **wJS** を提案する

$$wJS(L, R) = |L|D_{KL}(L||P) + |R|D_{KL}(R||P)$$

多様な解像度の補助空間データを用いた 空間データの高解像度化

田中佑典 (NTT) 岩田具治 (NTT) 田中利幸 (京大) 倉島健 (NTT)
大川真耶 (NTT) 戸田浩之 (NTT)

目的

補助データを活用して空間データを高解像度化し、都市計画等を利用

手法

補助データの空間補間における信頼度を組み込んだガウス過程モデル

結果

補助データの空間粒度の違いを効果的に活用した高精度な予測を実現

提案: 低解像度データ $\mathbf{a}$ の生成モデルを設計し、高解像度データ $\mathbf{z}$ をベイズ推定

(1) 補助データの予測分布を導出

$$p(\mathbf{f}_s^*) = \mathcal{N}(\mathbf{f}_s^* | \bar{\mathbf{f}}_s^*, \Sigma_s^*)$$

平均 共分散行列

(2) 高解像度データ $\mathbf{z}$ を線形回帰でモデル化

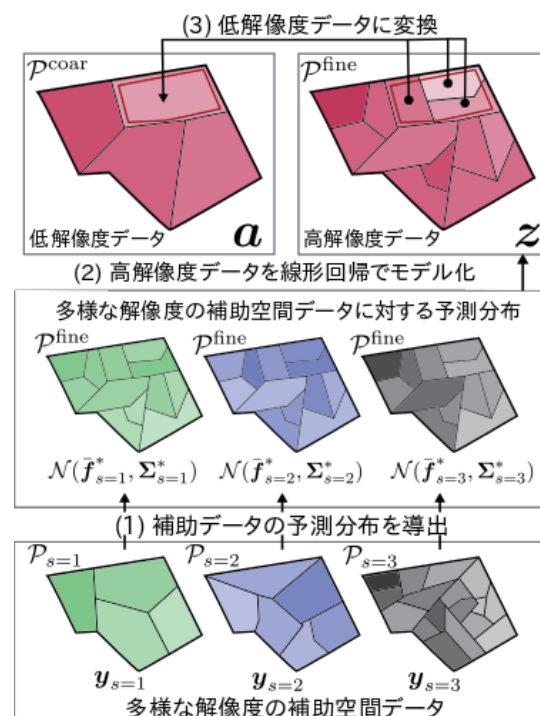
$$p(\mathbf{z} | \mathbf{F}^*) = \mathcal{N}(\mathbf{z} | \mathbf{F}^* \mathbf{w}, \mathbf{K})$$

$\mathbf{f}_s^*$ の行列 回帰係数 グラム行列

(3) 低解像度データ $\mathbf{a}$ に変換

$$p(\mathbf{a} | \mathbf{z}) = \mathcal{N}(\mathbf{a} | \mathbf{H}\mathbf{z}, \sigma^2 \mathbf{I})$$

変換行列 ノイズの分散



小型霊長類における動作のマーカレス測定と時空間モデリング

三村 喬生 松本 惇平 持橋 大地 島谷 健一郎 中村 友昭 長井 隆行
西条 寿男 須原 哲也 南本 敬史

目的 小型霊長類の社会的交流における身体表現の時系列構造を推定する

手法 3次元モーショントラッキング技術を新規開発し、特徴量時系列を分節化

結果 マーカレスに身体軌跡と顔の向き計測し繰り返し文脈の分節化に成功

- コモンマーモセットの自由運動から身体情報(姿勢・動作)を抽出するために、深度カメラ4台による計測情報に基づき、骨格モデルの当てはめ(力学演算)と顔検出(深層学習)を行い、マーカレスで身体部位(顔・頭・腰・尻)の空間位置を 3mm^3 グリッド, 30 frame/sec精度で推定した。
- 行動特徴量（床面からの頭高・体軸角度変化量・顔の向き変化量）を算出し、主成分分析による低次元化の後、ガウス過程隠れセミマルコフモデルによる分節構造推定を行った。
- 繰り返し餌取り行動文脈において、餌取りタイミングを適切に分節化する事に成功した。今後、より汎用性を高め、社会的行動とその障害モデルにおける身体表現の時系列構造解析に実装する予定である。

カーネル求積による積分変換の計算

園田 翔（理研AIP）

目的 数値積分を用いたニューラルネット学習法の近似誤差を評価したい

手法 カーネル求積を用いてニューラルネットの積分表現を数値積分します

結果 （重点サンプリングは強い）

• 数値積分によるニューラルネット学習法

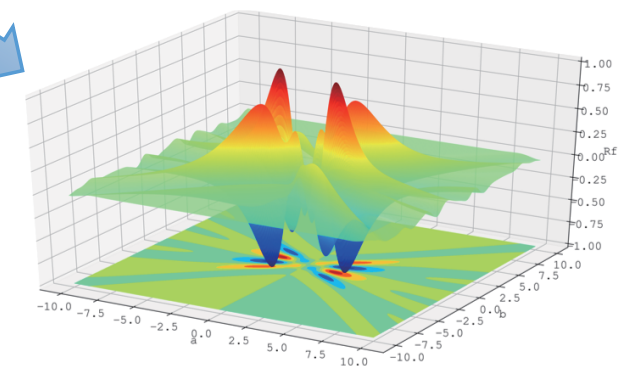
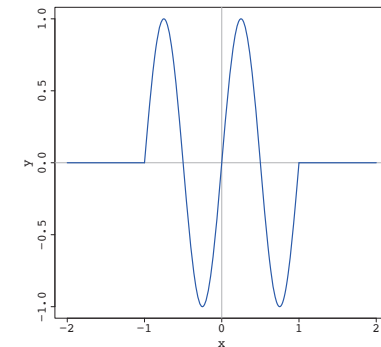
- In[0]: データ $y_i = f(x_i) + \varepsilon_i$
- Out[0]: リッジレット変換 $f^\sharp(a, b)$
- Out[1]: ニューラルネット（数値積分）

$$S[f^\sharp](x) = \int f^\sharp(a, b) \sigma(a \cdot x - b) da db$$

$$\approx g_p(x) = \sum_{j=1}^p c_j \sigma(a_j \cdot x - b_j)$$

• カーネル求積による近似誤差評価

$$\|S[f^\sharp] - g_p\|_\infty \lesssim 1/p \leftarrow \begin{array}{l} 1. \text{一様評価} \\ 2. \text{MC積分より速い} \end{array}$$



頑健な逐次グラフ型 半教師付学習アルゴリズムの提案

和田 裕一郎 (名大) Su Siqiang (香港理工大学)
熊谷 亘 (理研AIP) 金森 敬文 (東工大/理研AIP)

目的 高速かつ高精度な逐次グラフ型半教師学習アルゴリズムの提案

手法 頑健で高速なラベル推定手法を提案

結果 既存手法に対して精度・計算時間の両側面で優位性を確認

問題設定と目標

- データが逐次的に到着
- ラベル付きデータは稀に入手可能
- データ量は巨大
- ***高速かつ高精度**なラベル予測が必要

提案手法

- 1.データの大域構造を把握
- 2.ノイズ除去
- 3.ラベル推定

応用例



心電図自動監視



ネットワーク監視



車載カメラによる物体検知

結論

1.精度比較

	提案手法	既存手法
MNIST	0.679(0.021)	0.663(0.019)
USPS	0.700(0.020)	0.612(0.050)
optdigits	0.971(0.002)	0.961(0.031)
COIL	0.660(0.034)	0.592(0.062)
Yale	0.528(0.046)	0.471(0.048)
Vowel	0.969(0.004)	0.966(0.002)
UMNIST	0.637(0.028)	0.544(0.059)
ORL	0.722(0.029)	0.656(0.058)

2.計算量比較

$O(n^2)$ vs $O(n^3)$

提案

既存

Generative Adversarial Networkを用いた 複数モードを持つネットワークトラヒックの異常検知

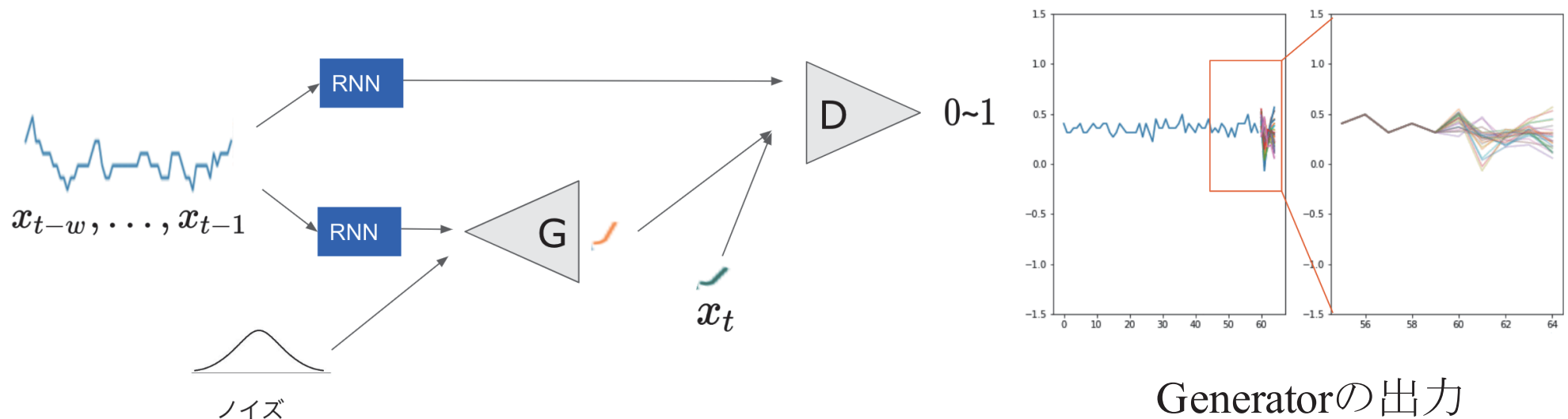
浅野 秀平 (NTTコミュニケーションズ) 大木 基至 (NTTコミュニケーションズ)

目的 複雑な揺らぎを持つネットワークトラヒックの異常検知

手法 異常度の計算にGANのDiscriminatorの出力を使用

結果 予測誤差に基づく手法と比較し高い精度での異常検知

- 異常度を予測誤差からでなく直接データから計算する手法として
Generative Adversarial Networkに基づく手法を提案
- 複数のモードを持つデータに対し、one-shot learningの考え方を導入



VGGモデルの視覚野的解釈における解釈の検討

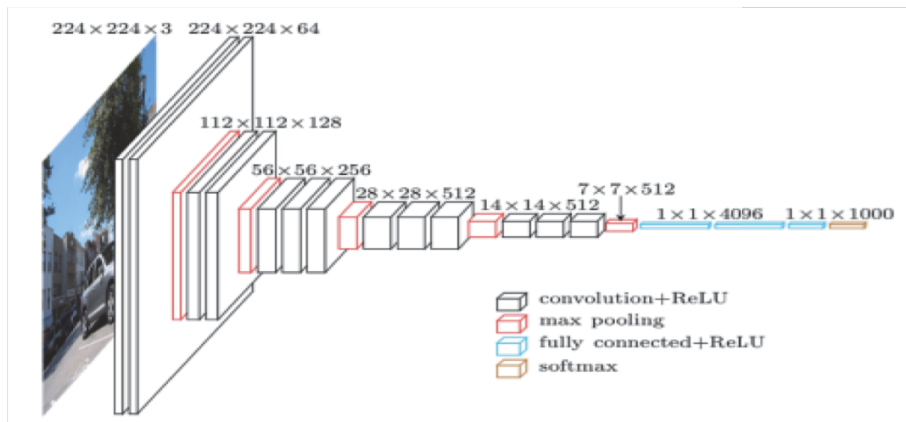
寺元 陶冶(電通大) 庄野 逸(電通大)

目的 視覚野的解釈を用いたVGG16モデルの解析

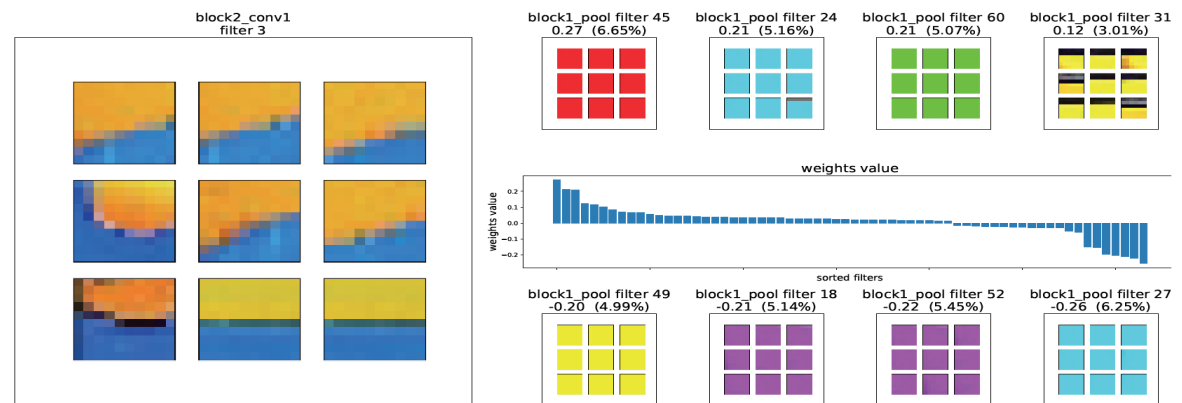
手法 VGG16が持つ重みと視覚野に見られる方位連続性との比較

結果 方位連続的なコラム構造は損失関数の最小化に寄与（かも）

- 低次の層から高次の層までの解析を行った。



VGG16 [K.Simonyan+,14]



四層目における受容野画像と、
色の方位連続性構造

乱流火災の時空構造の基礎的解明

富上 拓光 (東理大) 八條 貴誉 (東理大院) 後藤田 浩 (東理大)

目的

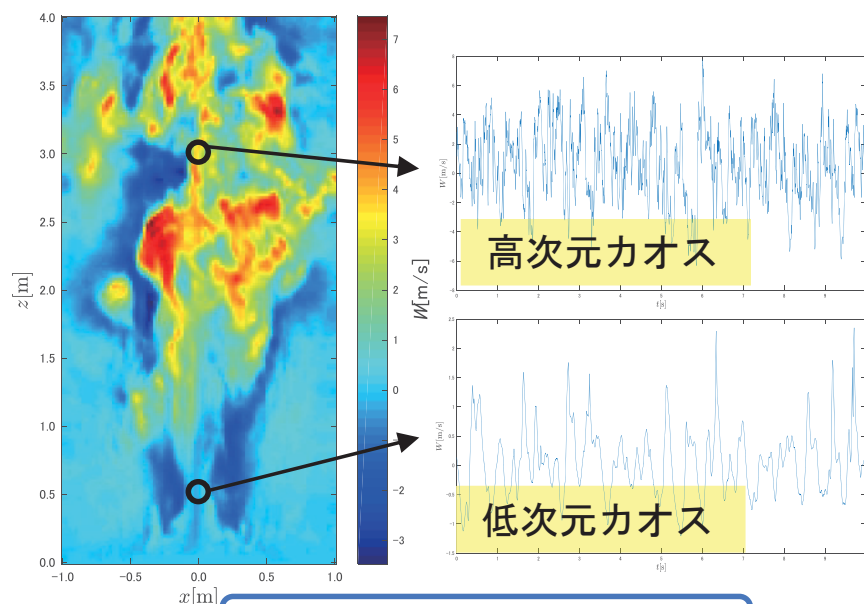
複雑系数理の基礎理論を用いた乱流火災の時空構造の解明と予測

手法

- 因果性解析 : Transfer Entropy
- Reservoir Computing : Echo State Network

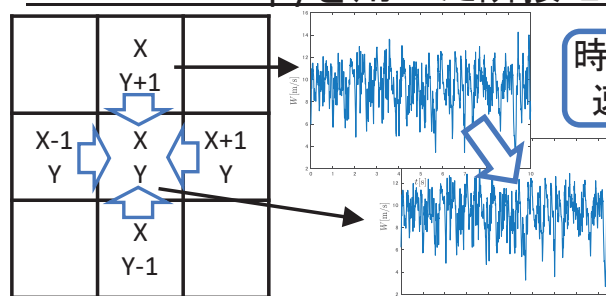
結果

- 変動の空間伝播は渦の分布と密接な関係がある
- 上流領域でのカオス的変動の予測可能性



乱流の発達⇒乱雑さの増大

Transfer Entropyを用いた隣接セル間の各方向因果性推定

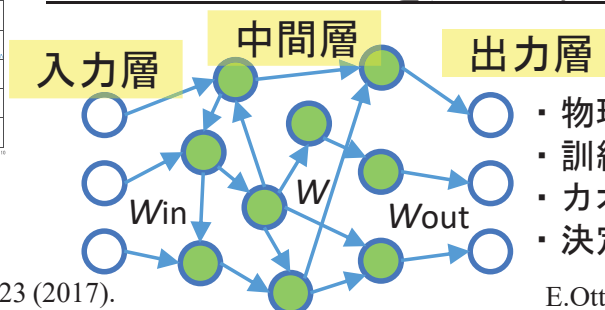


時系列の有向因果性解析手法を用いて
速度変動の空間的な伝播構造を解明

- Transfer Entropyを用いた解析
- 渦構造との関連を考察

T.Schreiber *et al.* Phys. Rev. L **85**, 461(2000).

Echo State Networkを用いた低次元カオス領域の速度変動予測



- 物理法則等のモデル化を必要としない予測
- 訓練により結合重み W_{out} のみを調整
- カオス的変動の予測可能性を考察
- 決定論性の強弱により予測精度が変化

ベイズ的変数選択法に基づく分光スペクトル分解

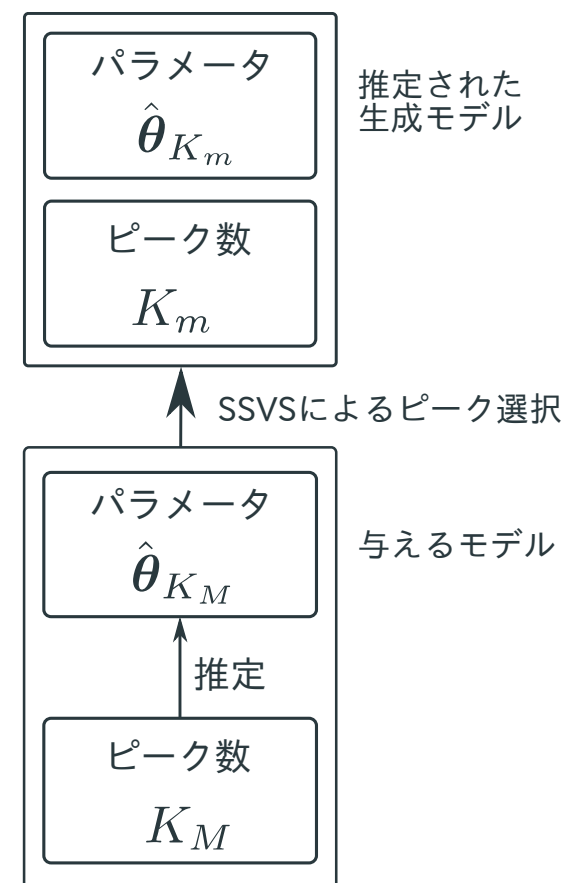
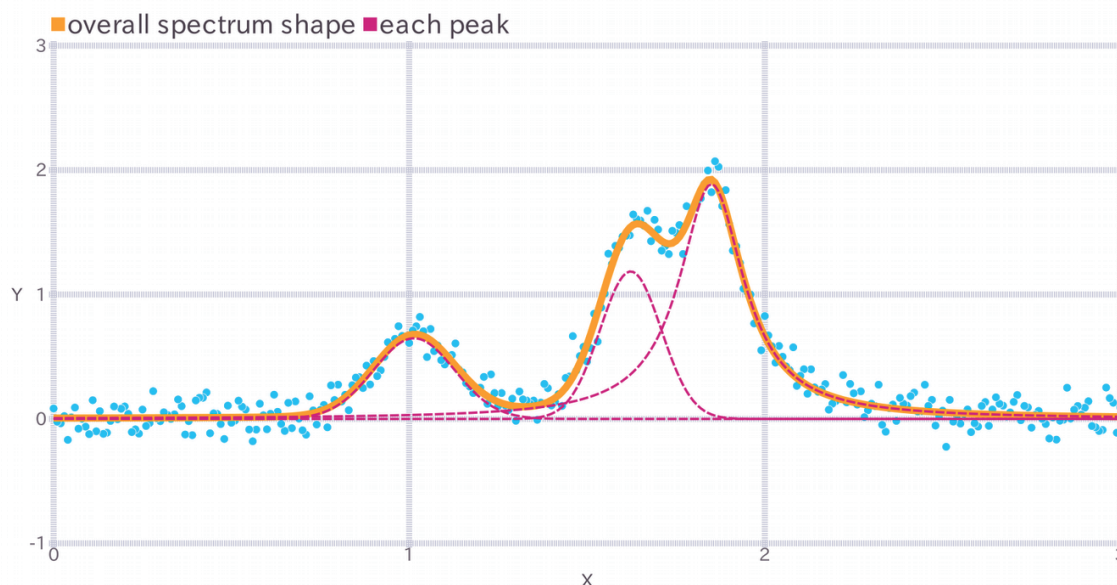
川島 貴大 (電通大) 庄野 逸 (電通大)

目的 分光スペクトルのピーク数とパラメータの同時推定

手法 spike and slab事前分布に基づくベイズ的変数選択法SSVSを適用

結果

- 人工データに対して有用性を確認
- 応用としてピーク形状の選択も可能



時系列を学習した潜在変数空間を持つVAEと 動画のフレーム補間への応用

下平 勇斗, 河合 新, 延原 肇 (筑波大)

目的

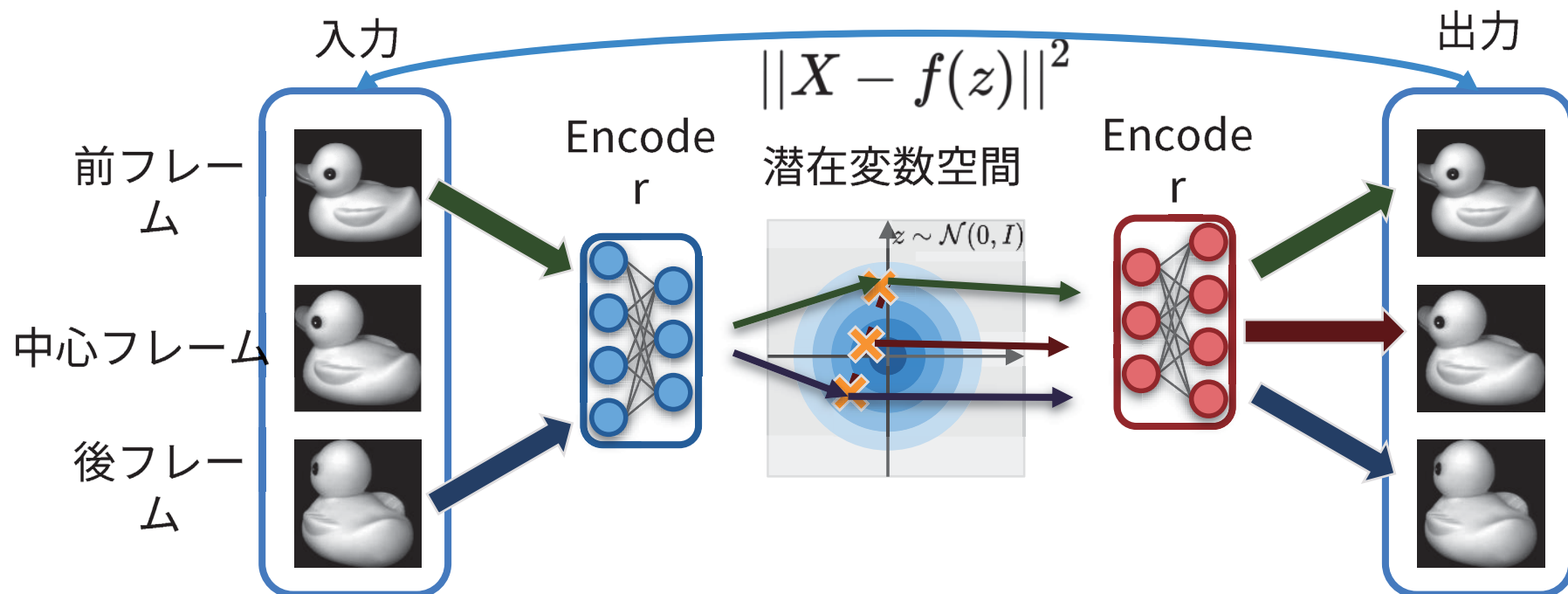
- 生成モデルに基づく動画のフレーム補間

手法

- 時系列画像を学習した潜在変数空間を持つVAEの提案
- 潜在変数空間上の補間、サンプリングによるフレーム補間の実現

結果

- 連続変化する画像の生成は成功



Monte-Carlo Search in Growing Single Rooted Directed Acyclic Graph by Best Arm Identification

Aurelien PELISSIER Atsuyoshi Nakamura Koji Tabata (Hokkaido University)

目的

手法

結果

履歴依存方策と定常方策の組み合わせによる 近似モデルベース・ベイジアン強化学習アプローチ

菱沼 徹（京大） 泉田 啓（京大）

目的 ベイズ適応的マルコフ決定過程 (BAMDP) の近似解

手法 定常方策と履歴依存方策を繰り返し交互にプランニングして組合せる

結果 数値実験を通じて有効性を確認

○ 背景・課題

- ・ BAMDPの最適解は、環境モデル不確定性の下での実サンプル効率を最大化する方策である
- ・ 最適解は履歴依存方策であるが、この空間での直接的なプランニングは難しい
- ・ 定常方策の空間でプランニングされたロバスト方策は、不確定性が大きいと有効に機能しない

○ 研究内容

- ・ 「時間が経過して不確定性が小さくなった後には、ロバスト定常方策が有効に機能する」と考える
- ・ ある時刻まで履歴依存方策、それ以降は定常方策を用いる、という方策の空間を扱う
- ・ 履歴依存方策と定常方策を繰り返し交互にプランニングして組合せて、近似解を構成する
- ・ 「地面摩擦係数が未知であるOpenAI Gym Hopperタスク」というBAMDPで、数値実験を行う

パスワイズ比較によるベイズ最適化

山下 拓真（東大） 小山 裕己（AIST） 佐藤 一誠（東大）

目的 探索と活用のトレードオフを行うスライダー空間の構築

手法 スライダーと探索空間の対応づけ（＝パス）の構築手法を新たに提案

結果 多くのベンチマーク関数において既存手法よりも早く最適解に収束

問題設定

スライダー入力の繰り返しによる最適化

例題：写真の色調補正パラメータの最適化

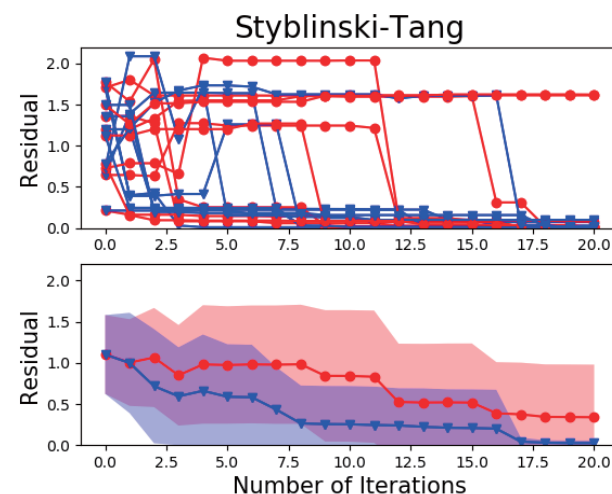
ユーザは提示されたスライダーを操作し最も良い状態を探索する



ベイズ最適化を利用し、スライダー位置と探索空間との対応づけを決定

実験結果

既存手法に比べて早く最適解に収束



全試行

平均

既存手法

提案手法

行列式強化学習

恐神 貴行 (IBM) ルディー レイモンド (IBM)

目的

強化学習によって複数のエージェントに協調した行動を選ばせたい

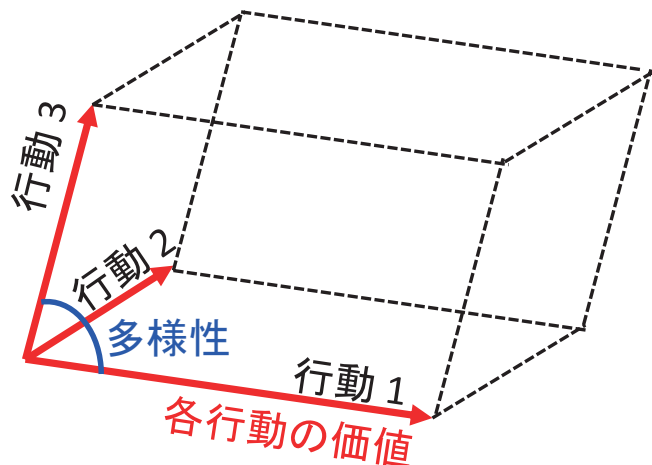
手法

価値関数を行列式を用いて表し、行動間の類似度を経験から学習する

結果

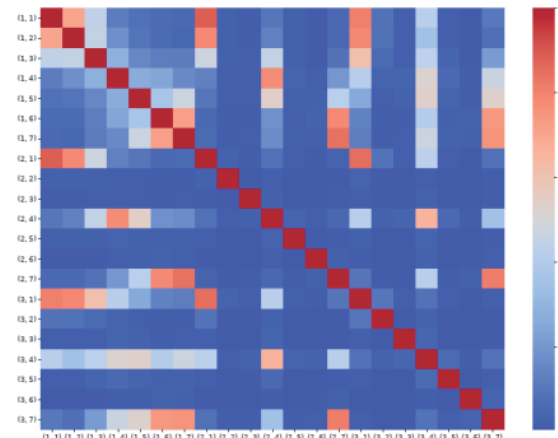
協調した行動が高い価値を持つように、類似度・価値関数が学習された

行列式を用いて
価値関数を近似

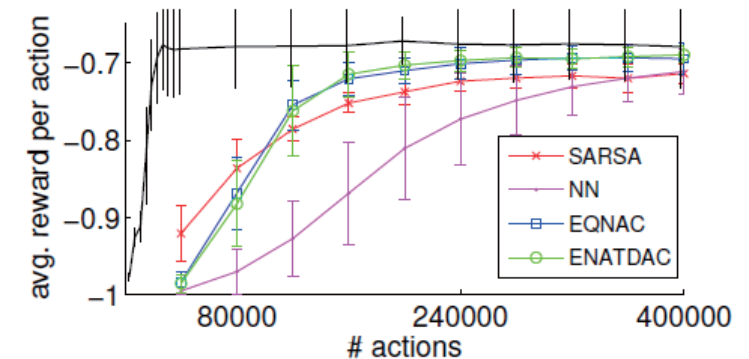


$$\text{体積} = \sqrt{\det(\mathbf{V}(\mathbf{x}) \mathbf{V}(\mathbf{x})^T)}$$

行動間の類似度
を経験から学習



複数エージェントの強化
学習で最高性能を達成



化学構造式のためのハイパーグラフ文法と分子の最適化への応用

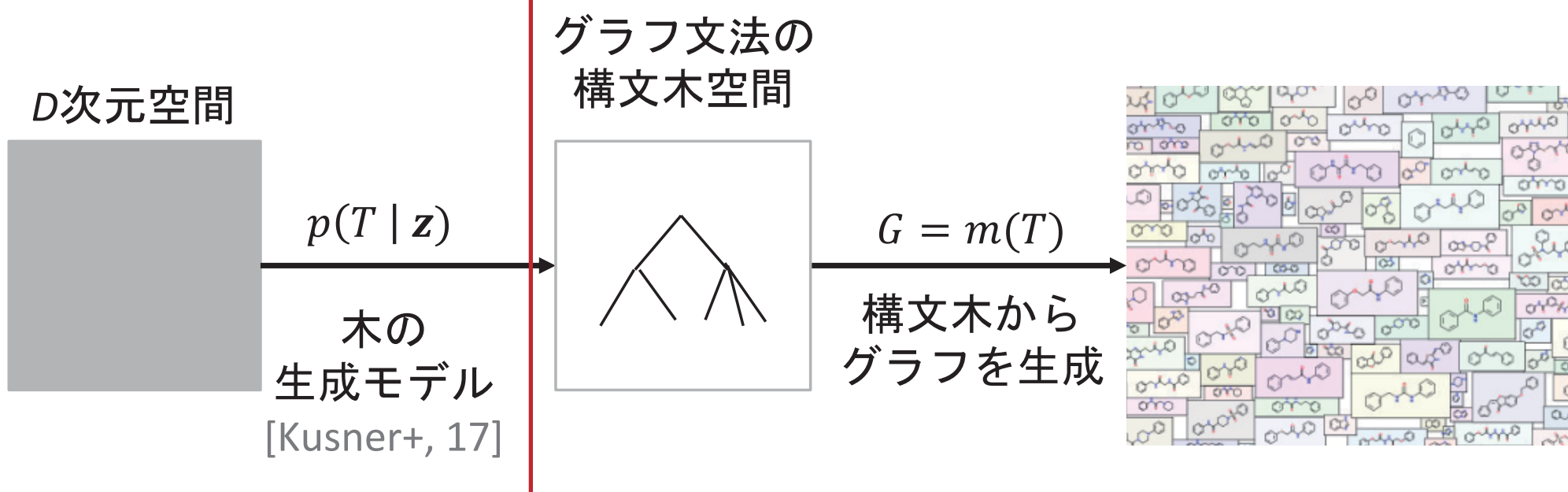
梶野 洸 (IBM Research - Tokyo)

目的 必ず原子価を守った構造式を生成できるDNNモデルを作りたい

手法 データから学習したグラフ文法を用いて構造式を生成

結果 生成される構造式は原子価を必ず守ることを証明
VAEを用いた分子最適化へ応用し、既存手法を上回る性能

提案手法



Selective Inference による画像セグメンテーション結果の統計的有意性検定

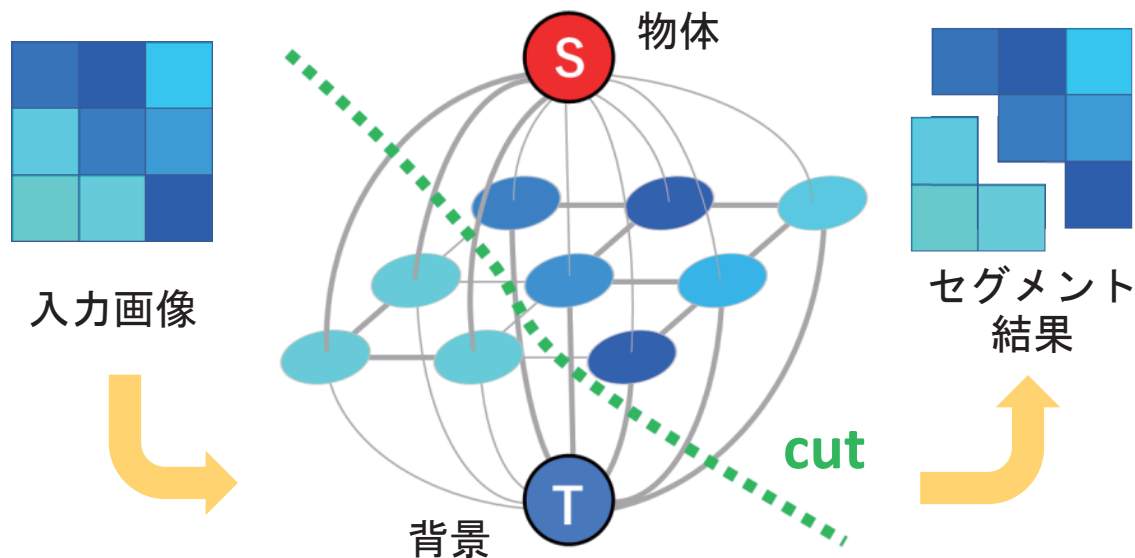
谷崎 光祐 (名工大) 稲津 佑 (理研) 橋本 典明 (名工大)
本谷 秀堅 (名工大) 竹内 一郎 (名工大・理研・NIMS)

目的 画像セグメンテーションによって得られた領域に関する仮説検定を行う

手法 アルゴリズムによってセグメンテーション結果が選択されたと考え、その**選択事象を条件づける**ことで、データ選択の影響を考慮する

結果 データ選択の影響を考慮したバイアスの無い非漸近的な検定を行えた

Step 1. グラフカットによる画像セグメンテーション



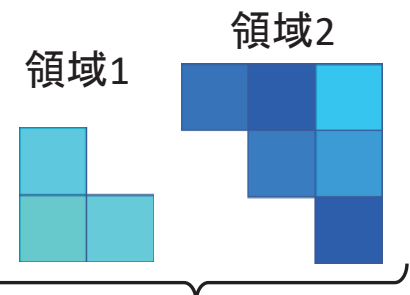
Step 2. 統計的仮説検定

得られた領域の母平均に関する二標本検定

$$H_0: \mu_1 = \mu_2$$

vs.

$$H_1: \mu_1 \neq \mu_2$$



$$p = 0.2785$$

セグメント結果を考慮した
選択的なp値を算出

温度付きsoftmax関数を含む深層学習器の実験的性能評価

岡田 莉奈、長谷川 聡 (NTTセキュアプラットフォーム研究所)

目的

分類タスク用の温度付きsoftmax関数の性能(正解率)を実験評価すること。

手法

通常の学習の場合とmodel extraction attackへ応用した場合の正解率を比較。

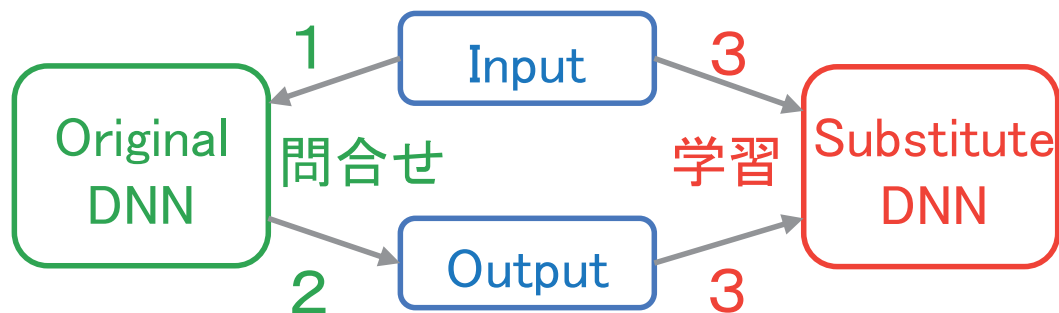
結果

通常の学習では性能が出ず、model extraction attackでは性能が出た。

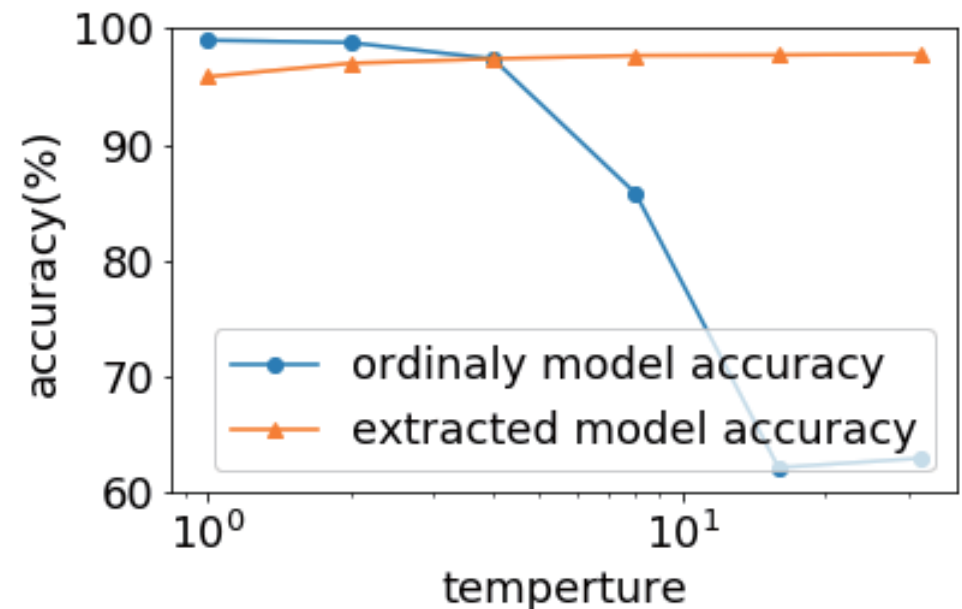
温度付きsoftmax関数

$$y_i = \text{softmax}(x, i, T) = \frac{\exp(x_i/T)}{\sum_{j=0}^{|x|} \exp(x_j/T)}$$

Model Extraction Attack



実験結果 (一例: MNIST)



貪欲法によるマルチタスク特徴選択

藤井海斗（東京大学）

目的 マルチタスク特徴選択に対して高速な近似アルゴリズムを提案

手法 二段階の組合せ最適化問題として定式化し、貪欲法を適用

結果 提案法は既存法と同じ近似保証でより高速、かつ実験的にも高性能

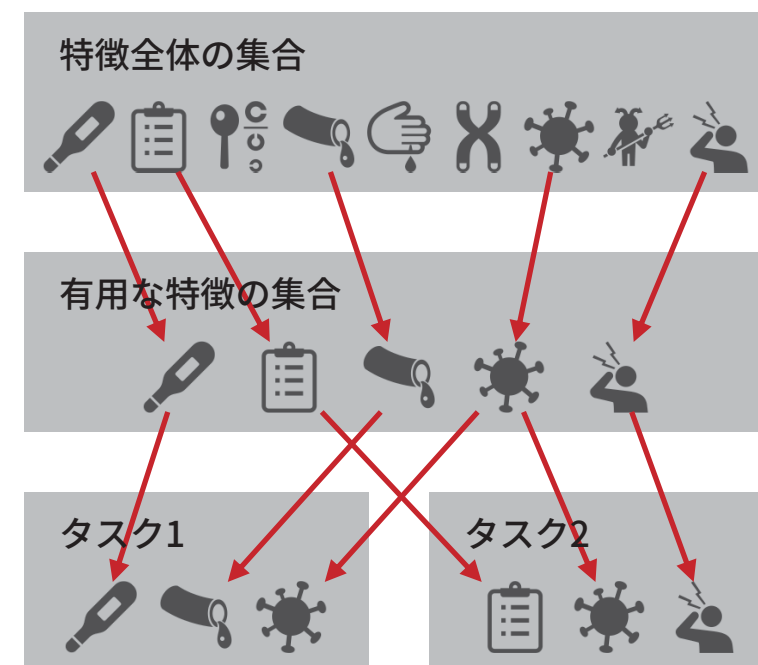
- マルチタスク特徴選択を二段階の組合せ最適化問題として定式化：

特徴全体の集合から有用な特徴を抽出し、
そこからさらに各タスクに用いる特徴を選択

- Replacement Deletion-OMPを提案：

置換貪欲法にOMPのアイデアを導入

- パラメタの復元に関する保証
- より一般的な疎性制約への拡張



反復型ラベルなし・ラベルなし分類による ラベルなしデータのみからの二値分類

梶 洋隆(トヨタ自動車), 杉山 将(理研AIP・東大)

目的

一つのラベルなしデータセットのみから二値分類を実施

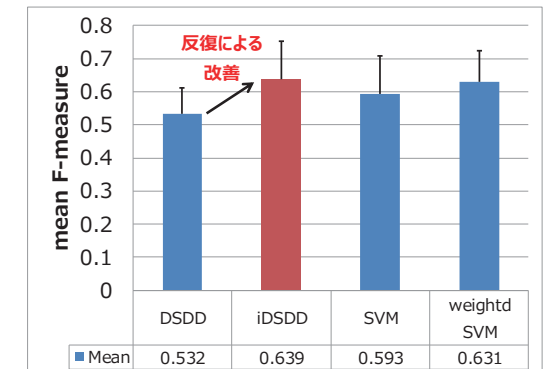
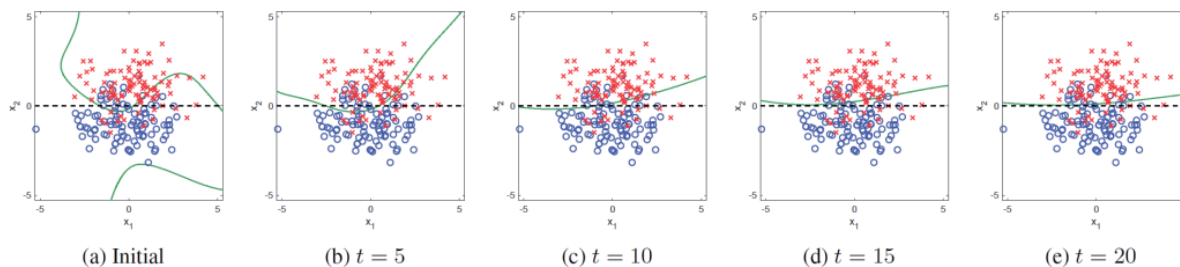
手法

ラベルなしデータを2分割, ラベルなし・ラベルなし(UU)分類を反復的に適用

結果

数値実験と実問題(眠気予測)で良好な性能を確認

- 分類問題ではラベル付けのコストが課題
- UU分類: クラス事前確率の異なる2セットのラベルなしデータから二値分類器を学習
- 本研究: 1セットのラベルなしデータから近似的にUU分類を行う手法を提案
- アイデア: k -means等で2分割したセットを初期値とし, UU分類を反復して性能改善



- 心拍情報からの眠気予測問題に適用 ⇒ 教師あり学習(SVM)と遜色ない結果

Softmax bottleneckの再考

金井 関利(NTT SIC・慶大) 藤原 靖宏(NTT SIC)
山中 友貴(NTT SC研) 足立 修一(慶大)

目的

Softmax bottleneckを解消する新たな出力関数の提案

手法

Log softmaxの値域からsoftmax bottleneckを再解釈し原因を説明.

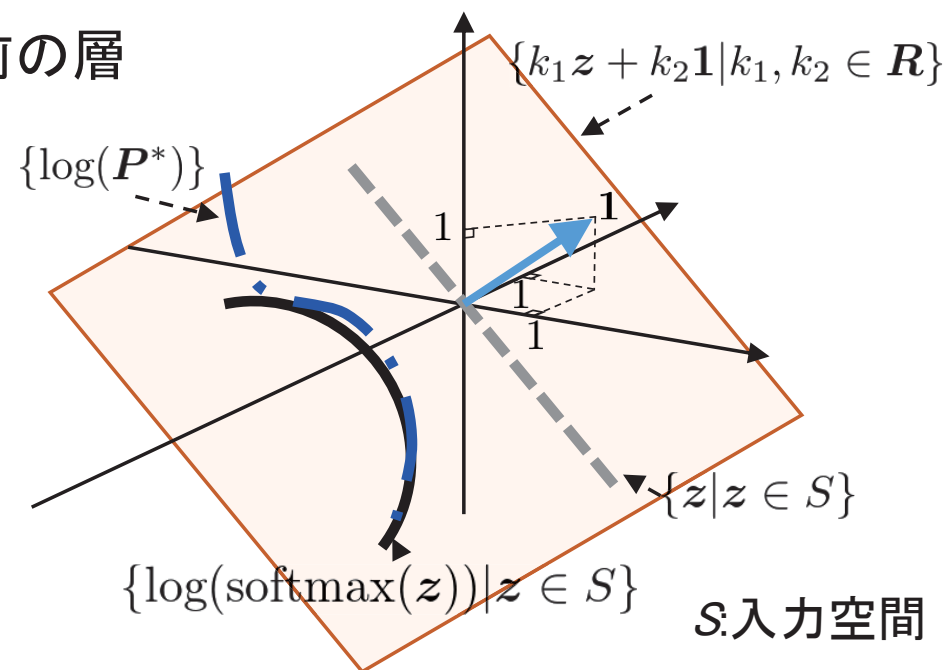
結果

指数関数だけを使うことが原因→sigmoidをかけることで解消, 精度改善

- Softmax bottleneckとは
 - Softmaxを使ったモデルの表現力が出力前の層のユニット数に依存してボトルネックとなる

- Log softmaxの値域から原因を説明
 - Log softmaxの値域は入力空間の次元+1次元の線形空間の部分集合

- ボトルネックを解消するための出力を提案
 - Sigsoftmax (sigmoid × exponential)
 - ReLU, sigmoidを使った場合も比較



Log softmaxの値域と真の確率 P^* の対数

ベイズ最適化を用いた高感度計測のための ESIパラメータ最適化法の検討

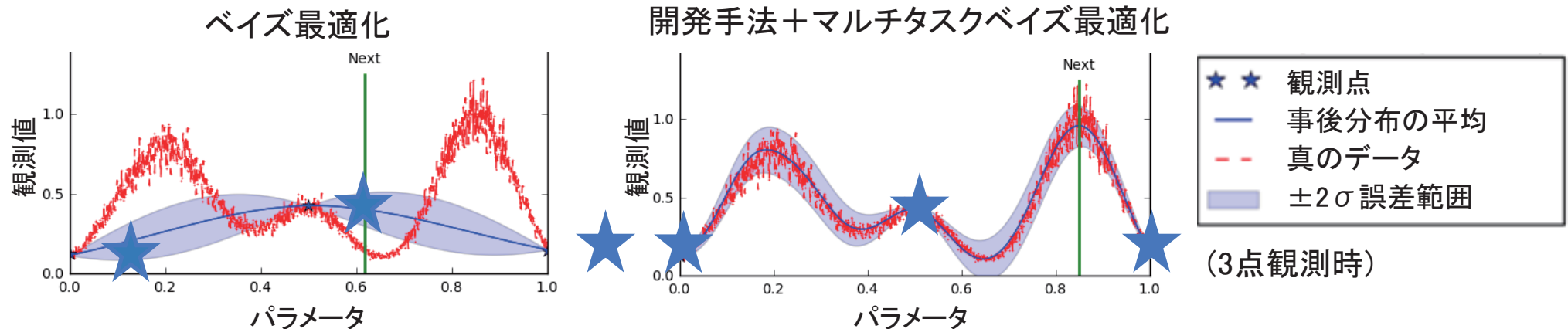
石川 勇樹 田川 雄介 Levi Mikael 森田 尚孝 渡邊 淳(島津製作所)

目的 高感度計測のために各種装置パラメータを最適化したい

手法 観測ノイズに対してロバストなマルチタスクベイズ最適化法を適用

結果 網羅計測に比べて30倍程度効率探索できることを確認した

- 装置の計測精度を考慮したReinterpolation処理(*1)により、観測ノイズに対してロバストな探索を実現
※1 Forrester et al., 2006
- マルチタスクベイズ最適化(*2)において、モデルの事前情報として類似の計測結果を用い、モデルのパラメータ空間でのシフト変化を計測結果から逐次補正する手法を開発
※2 Swersky et al., 2013



リーマン多様体上における 多変量低ランク回帰モデルの推定

吉川 剛平 (電通大) 川野 秀一 (電通大)

目的 多変量低ランク回帰モデルの推定アルゴリズムの改善

手法 リーマン多様体上での最適化アルゴリズム

結果 推定精度の向上・安定性が数値実験を通して示された

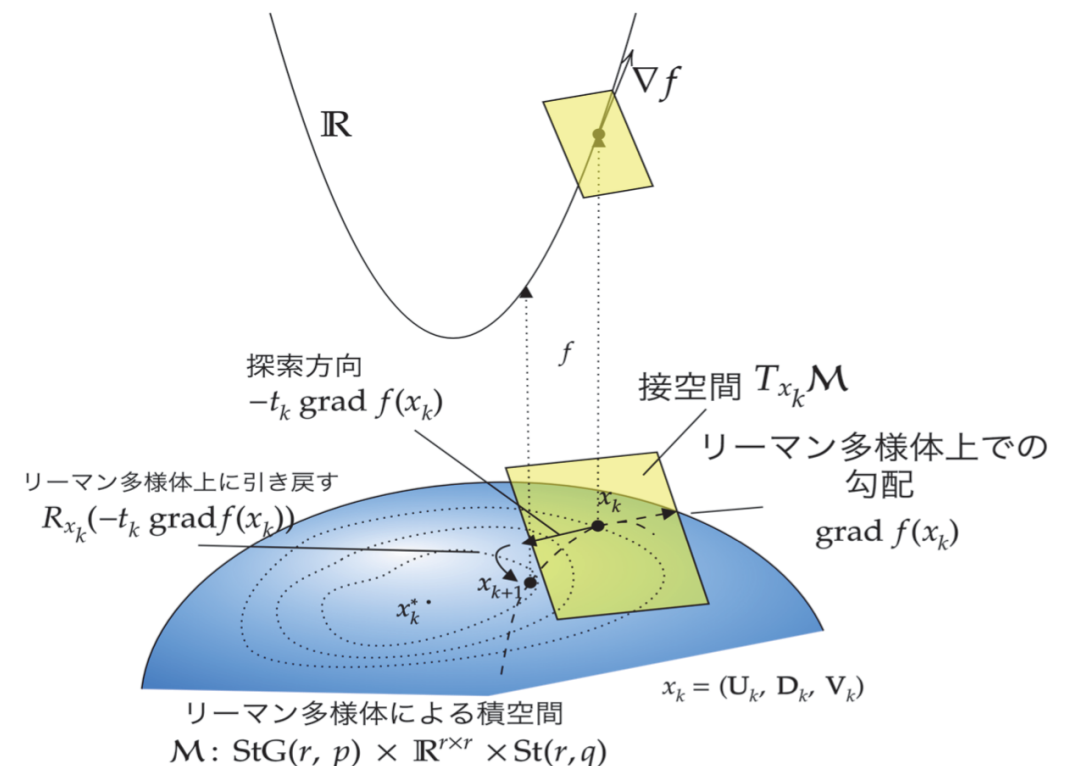
- 多変量回帰モデルの変数選択問題
 - 低ランク制約＋スパース制約
 - ✓ 大量の説明変数から **有用な情報を抽出**
 - ✓ モデルの **解釈性**



従来手法の推定方法は不安定

- リーマン多様体上での最適化問題
 - パラメータ空間が **リーマン多様体** を形成
 - 従来手法よりも **安定した推定** が可能

- リーマン多様体上での最適化の概要図



セーフスクリーニングによる最適ルールフィットモデルの学習

加藤 宏樹 (名工大) 花田 博幸 (名工大) 竹内 一郎 (名工大・理研・NIMS)

目的

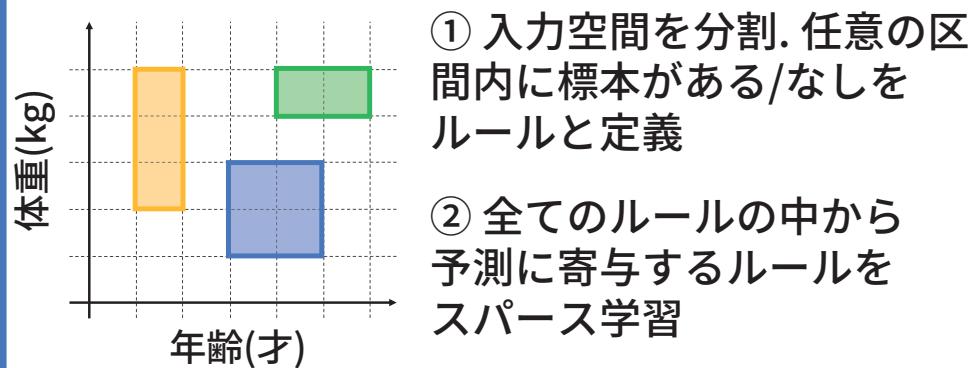
“ルール”のスパースな線型結合による予測モデルの学習を, 考えられ得る **全ての**ルールを対象に行う。

手法

ルールを木構造で表現し, 提案するセーフスクリーニング法で学習

結果

得られる解の**最適性を保証**しつつ, **効率的な学習**を実現

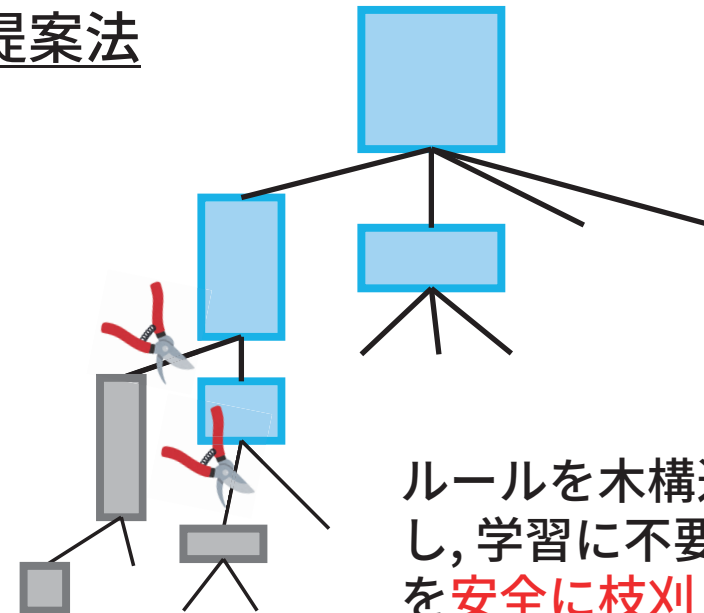


$$\hat{y} = w_1 \text{ (orange box)} + w_2 \text{ (green box)} + w_3 \text{ (blue box)}$$

提案法

③ 獲得されたルールで予測

提案法



高次元L1-ball上でMinimax Regretを達成する新しい事前分布

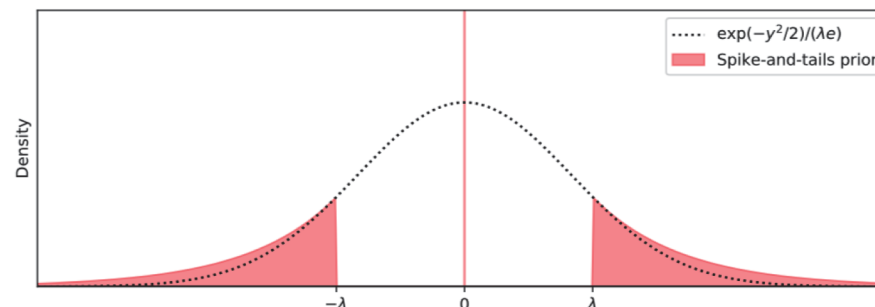
宮口 航平 (東大) 山西 健司 (東大)

目的 高次元確率モデルに対する最大リグレットを最小化する予測

手法 新しい事前分布を用いたベイズ予測分布を構成

結果 Adaptive Minimax 性の保証

- 背景) ベイズ予測分布を用いてmax regretを最小化したい
- 問題) どのような事前分布が良いか?
- 既存研究) “モデル次元 $\ll$ サンプル数” のとき $\rightarrow$ Jeffreys priorが最適
- 本研究) “モデル次元 $\gg$ サンプル数” のとき $\rightarrow$ Spike-and-Tails prior が最適[†]



Spike-and-Tails prior

[†]ユニバーサル定数を除いて

双曲空間における測地線更新アルゴリズム

榎田洋介(東大) 鈴木惇(東大) 山西健司(東大)

目的

双曲空間上での最適化アルゴリズムの開発と性質の解明

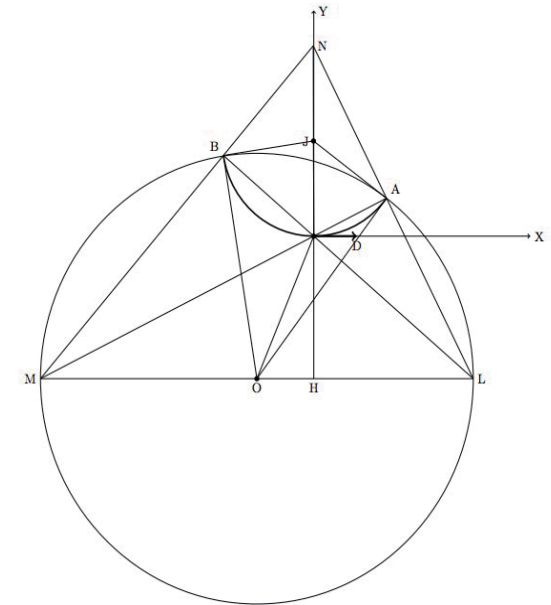
手法

測地線に沿った更新式の導出(初等幾何的)
収束レートの解析, 「自然勾配法」との比較

結果

測地線更新アルゴリズムは重心問題で一次収束
「自然勾配法」が最適点に収束しない例を構成

- 双曲空間...爆発的な広がりを持つ空間
 - 木構造やグラフを埋め込むための空間として注目される
- 測地線(最短線)に沿う最適化アルゴリズムの更新式を導出し, アルゴリズムの収束レートを導出
 - 測地線更新式の導出は初等幾何的(右上図参照)
 - 距離の2乗和を最小化する問題で一次収束(指数的に収束)
- 自然勾配法と比較
 - 自然勾配法は中点を求める問題において最適点に収束しない



確率的勾配降下法の収束性について

川島 孝行 (総研大) 藤澤 洋徳 (統数研)

目的 非凸損失関数・非凸正則化項からなる最適化問題のSGDによる収束性

手法 Stochastic Gradient Descent

結果 収束性は証明できた！しかし、収束レートに次元 p に依存した定数が出現...

設定

$$\cdot \text{主問題 } \arg \min_{\theta} f(\theta) + \sum_{j=1}^p \lambda_j h_j(\theta_j) \quad \cdot \text{更新式 } \theta_j^{(t+1)} = \text{prox}_{\eta_t \lambda_j h_j}(\theta_j^{(t)} - \eta_t G_j(x; \theta^{(t)}))$$

仮定

損失関数 f は非凸かつ微分可, 勾配がリプシッツ連続

正則化項 h_j は (下半連続な) 準凸 例) L_0 正則化, MCP, SCAD, Capped- L_1 , Log-sum... $+ \alpha$

比較

既存研究 (Xu and Xin 2016, Ghadimi and Lan 2016, Davis et al. 2018) は,
正則化項は凸またはリプシッツ連続な非凸 $\rightarrow L_0$ は含まない！

準凸にしたから p が出現？ (おそらく) 違うはず！ $\rightarrow$ 変数の分離の仕方によっては出ない！

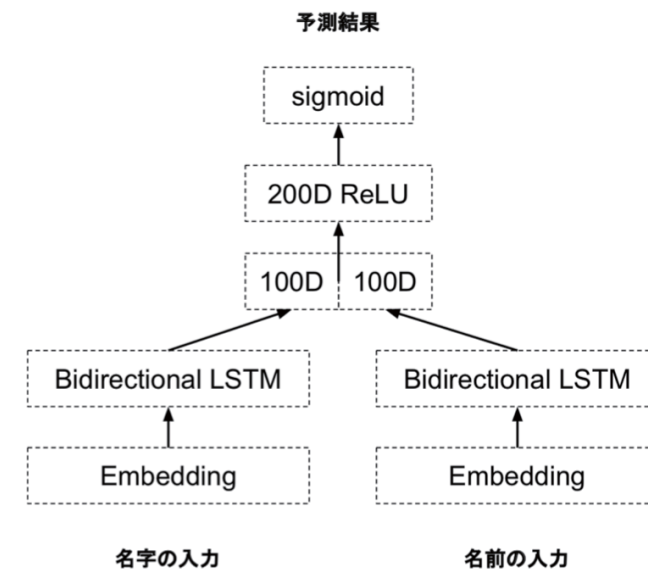
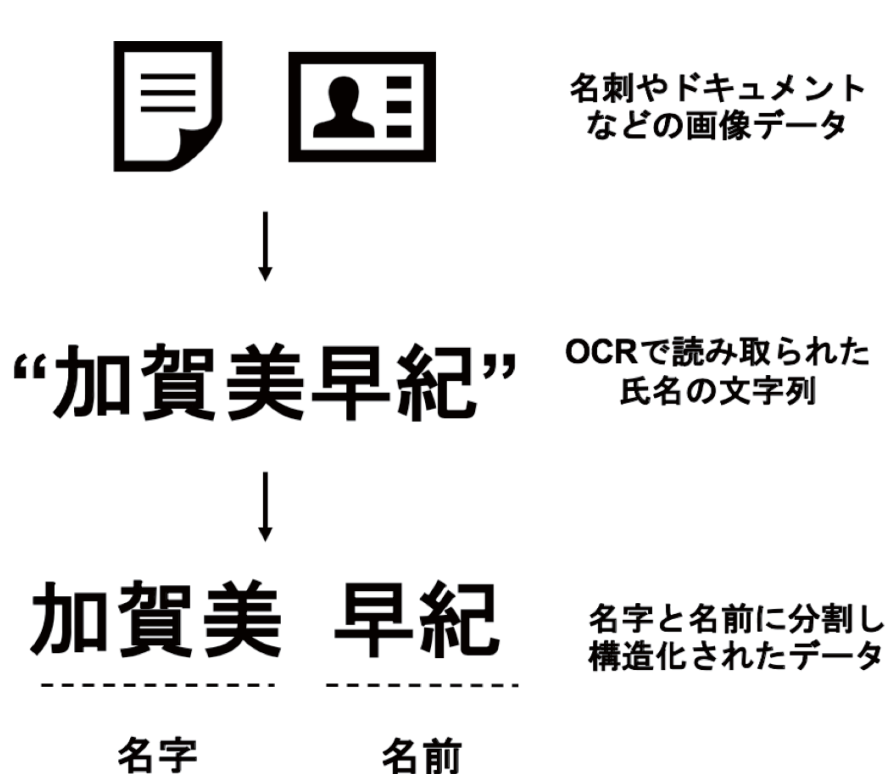
Bidirectional LSTMを用いた日本語氏名の姓名分割

奥田裕樹 高橋寛治 (Sansan株式会社)

目的 日本語氏名の分割手法の提案

手法 Bidirectional LSTMを用いて名字と名前を分割する位置を推定

結果 実在する氏名および自動生成された氏名ともに、正確度99%以上を実現



提案手法：姓名の分割候補の正しさを判定するBidirectional LSTMを構築

Statistical Logical Relationship Analysis using Tarone's Multiple Testing Correction

福永 津嵩（東大） 岩崎 渉（東大）

目的 統計的に有意に頻出する、アイテムの三項間論理関係を発見する。

手法 Taroneの多重検定補正法を用いる事で、高い統計的検出力を得る。

結果 微生物ゲノムのデータから、光合成に関する新規三項間論理関係を発見。

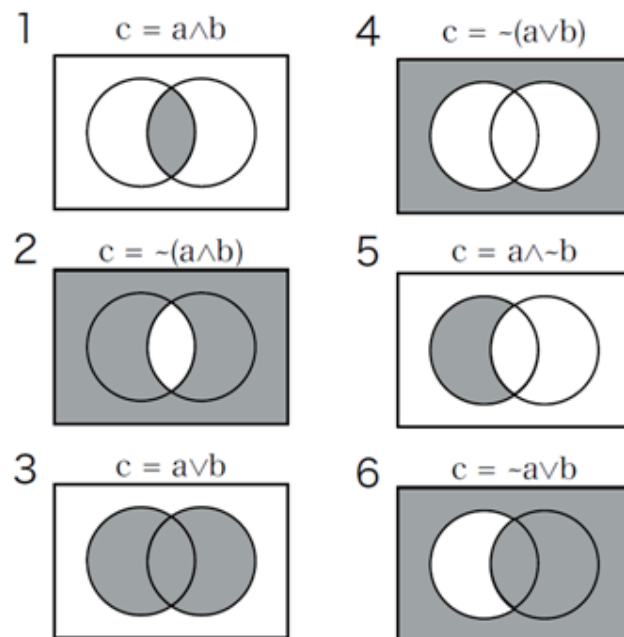


図1 三項間論理関係を示したベン図

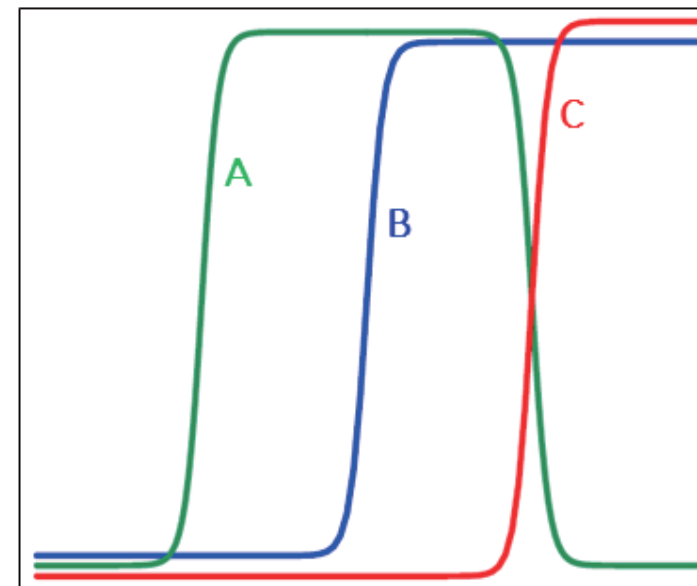


図2 $C = \neg A \wedge B$ を満たす遺伝子間関係のグラフ

Attention機構を用いた集合データに対する 確率的ニューラル回帰モデル（による共同研究者推薦）

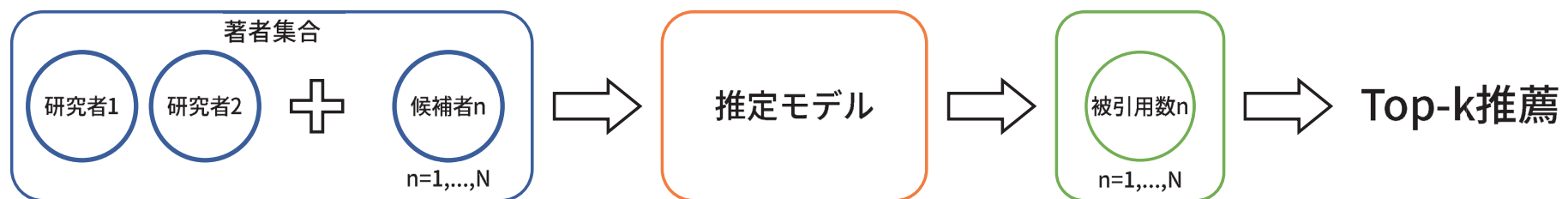
武田 悠佑(奈良先端大) 岩田 具治 (NTT) 澤田 宏 (NTT)

目的 より多く引用されるような研究を実現するための共同研究者推薦

手法 研究者集合を入力として回帰モデルで被引用数を推定しランキング

結果 提案モデルは現状まだベースラインと同程度の性能

- 共同研究者推薦：研究者に対する効果的な共同研究者の推薦
 - 従来：共著ネットワークにおけるリンク予測問題
 - 本研究：著者集合を入力とする論文の被引用数の予測問題
- 提案モデル：Deep Averaging Networks (DAN) +事前確率+Attention機構
 - 集合を入力として扱うため可変長の入力に対応できるDANを使用
 - 研究者同士の相互作用を考慮するためにAttention機構を導入



Long Short-Term Memoryアンサンブルによる 学習曲線の予測

尾崎 嘉彦 (産総研・グリー) 大西 正輝 (産総研)

目的

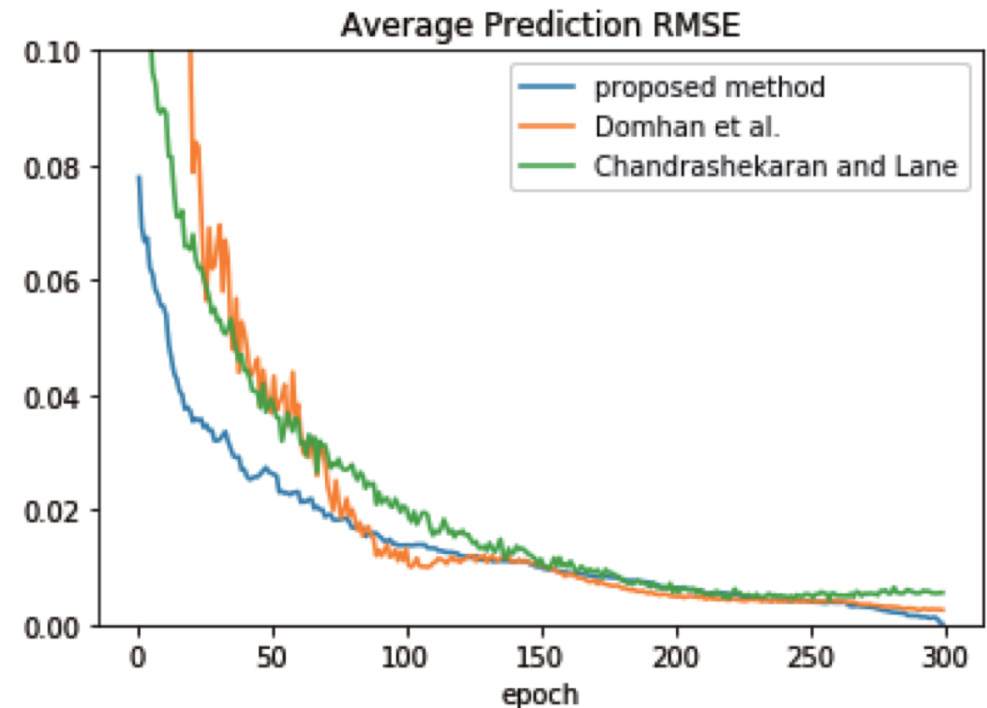
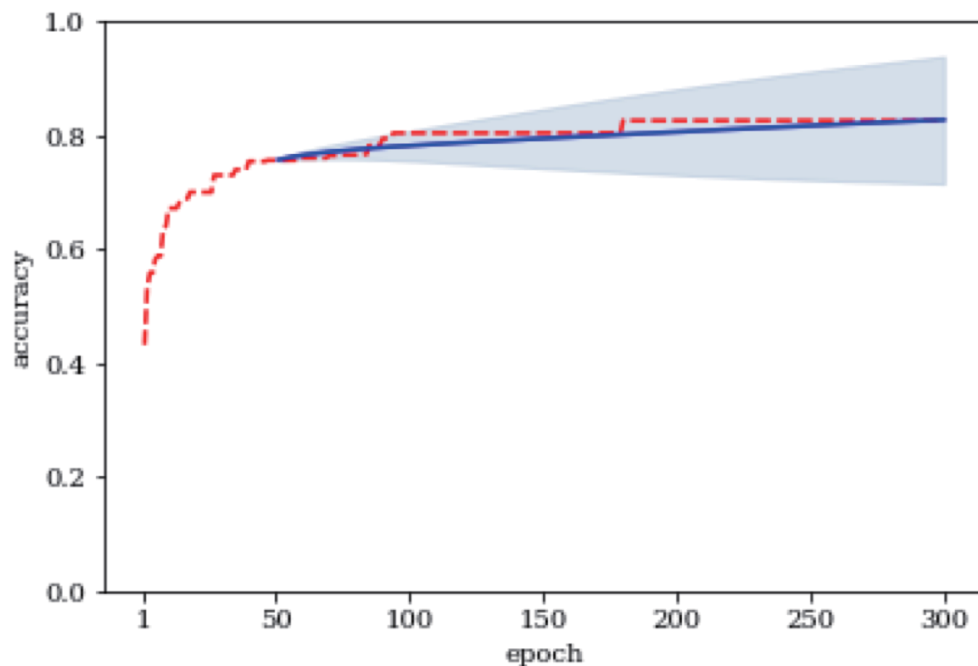
学習の途中経過から最終的な機械学習モデルの性能を予測

手法

過去の学習曲線を用いて訓練したLSTMをアンサンブルし予測曲線を生成

結果

従来手法に比べ高速かつ高精度な予測を達成



Canonical Time Warpingにおける勾配法によるパラメータ最適化手法の提案

河野圭祐, 小出智士, 沓名拓郎 (豊田中央研究所)

目的

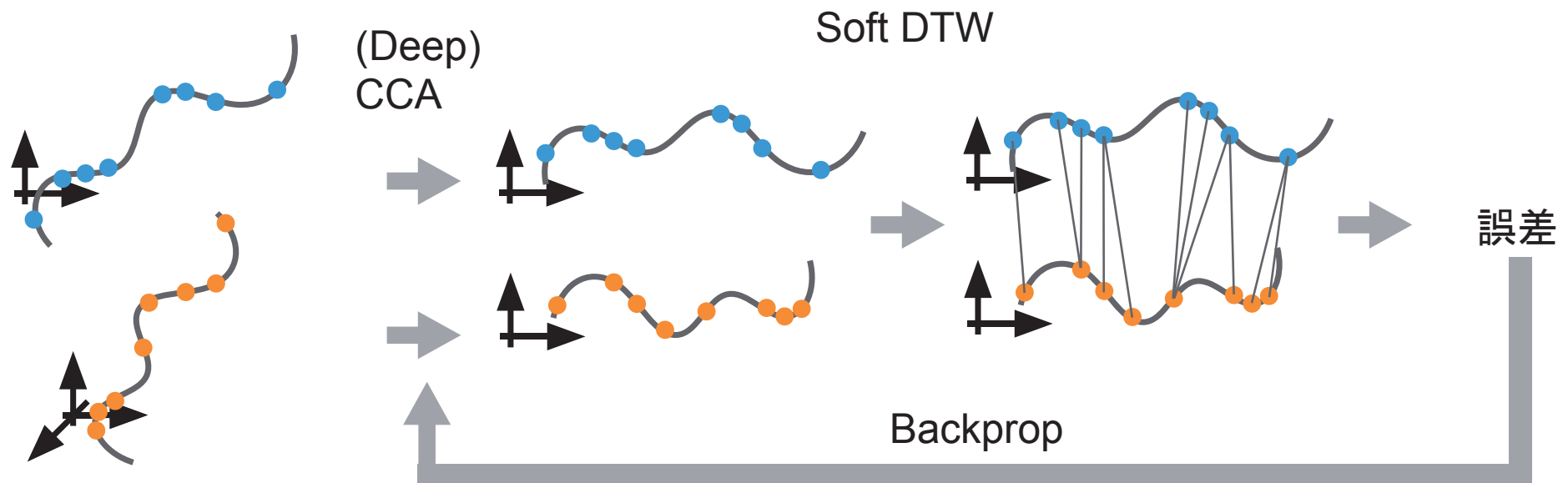
複数の多変量系列データに対する, 時間方向の伸縮等を考慮した, フレーム合わせ

手法

DCTW (DNN+CCA+DTW)を微分可能に拡張

結果

DCTWのend-to-endの勾配法による学習が可能になった.



分布間距離を用いたXNOR-Netの学習手法の改善

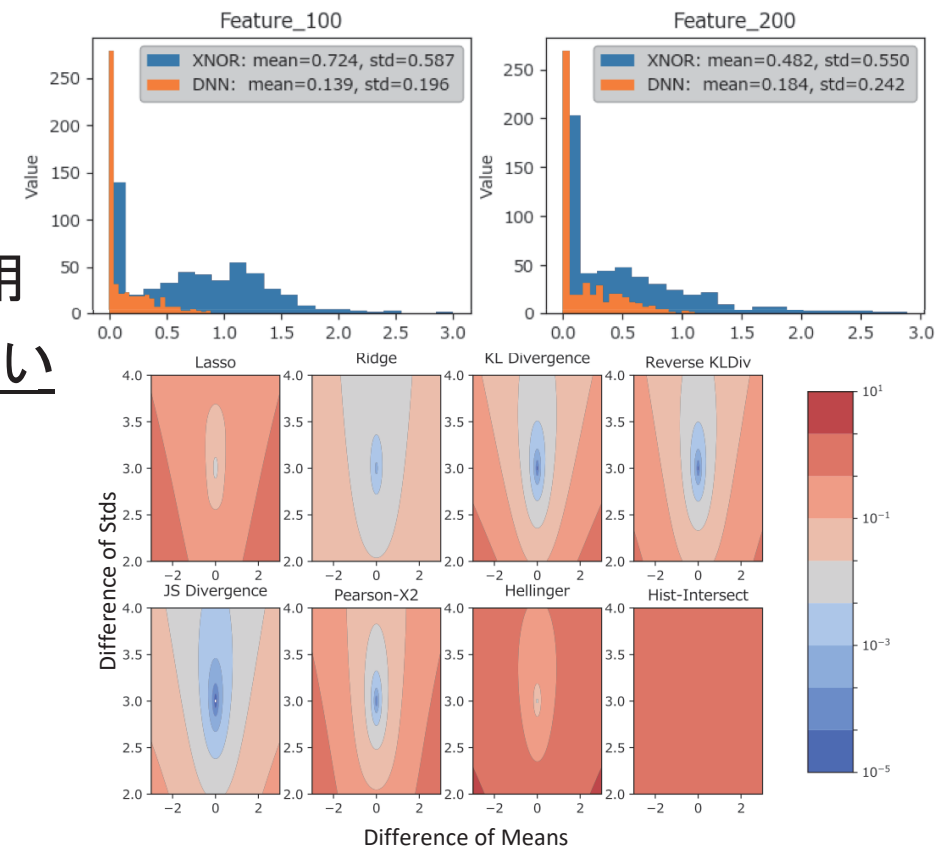
河野 慎 (慶應) 大屋 優 (NTT) 八木 哲志 (NTT) 中澤 仁 (慶應)

目的 ニューラルネットワークの二値化における性能の向上

手法 DNNとXNOR-Net間における最終層手前の特徴量の距離を最小化

結果 様々な距離尺度における挙動をもとに、XNORの性能検証を行なった

- 学習済みモデルの各特徴量の分布
 - 平均, 標準偏差共にXNOR>DNN
 - 先行研究[Xu et al., 2018]: L2ノルムを利用
 - L2では分散方向の広がりを許容しやすい
- XNORがDNNに近づきにくくなる
- 実験: L2以外の距離で精度検証
 - 分類問題に適用(CINIC-10)



周辺尤度を用いた ベイジアンネットワーク分類器のアンサンブル法

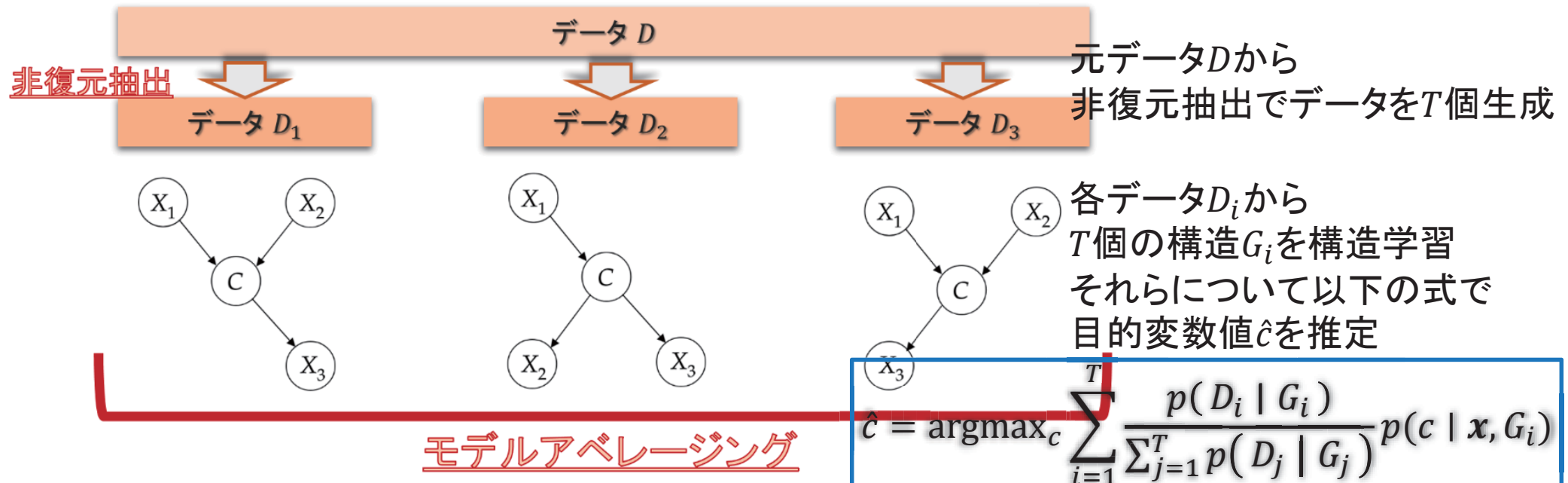
青見 樹（電通大） 植野 真臣（電通大）

目的 ベイジアンネットワーク分類器の精度の向上

手法 ベイジアンモデルアベレージングを用いて、分類器にアンサンブル法を適用する

結果 データサンプル数が十分でないデータに対し、分類精度が向上した

- ベイジアンモデルアベレージングを用いるにあたり、全ての候補構造の計算は計算時間が膨大
→ アンサンブル法で得られた構造についてのみ計算する



グラフ上のNP-Hard問題における難しいインスタンスの生成

佐藤竜馬（京大） 鹿島久嗣（京大）

目的 グラフ問題を解くアルゴリズムにとって時間がかかるグラフを生成する

手法 ニューラルネットワークと REINFORCE アルゴリズムで難しさを最適化

結果 いくつかの問題とアルゴリズムにおいて極度に難しいグラフを作れた

- 組合せ問題を解くアルゴリズムにとって解くのに時間がかかるインスタンスを発見することは、ベンチマーク生成やアルゴリズム解析にとって重要
- 既存の研究では、問題の性質を吟味することで難しいインスタンスの生成アルゴリズムを考案してきた
- 本研究では、このタスクを最適化問題として捉え、グラフ生成モデルと強化学習手法を用いて自動的に難しいインスタンスを生成する手法を提案
- 3-coloring, max clique, minimum vertex cover などの重要な問題で実験的に手法の有効性を検証した

潜在ドメインベクトルを用いたゼロショットドメイン適応

熊谷 充敏 (NTT) 岩田 具治 (NTT)

目的

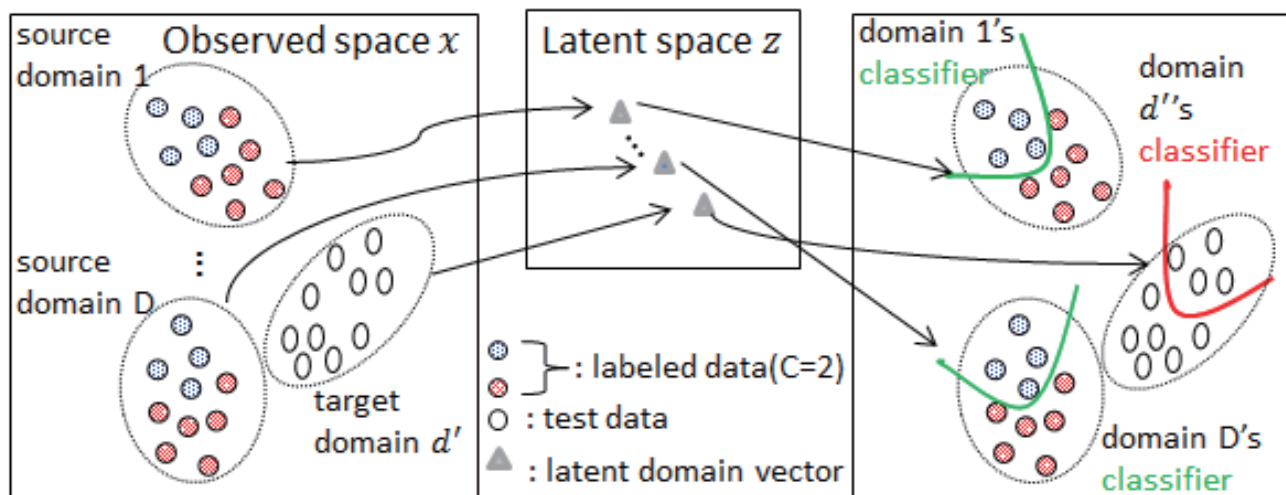
学習時に目標ドメインのデータが得られない、ドメインを特徴づける付加情報も得られない状況であっても、目標ドメインに適した教師ありモデルを推論したい

手法

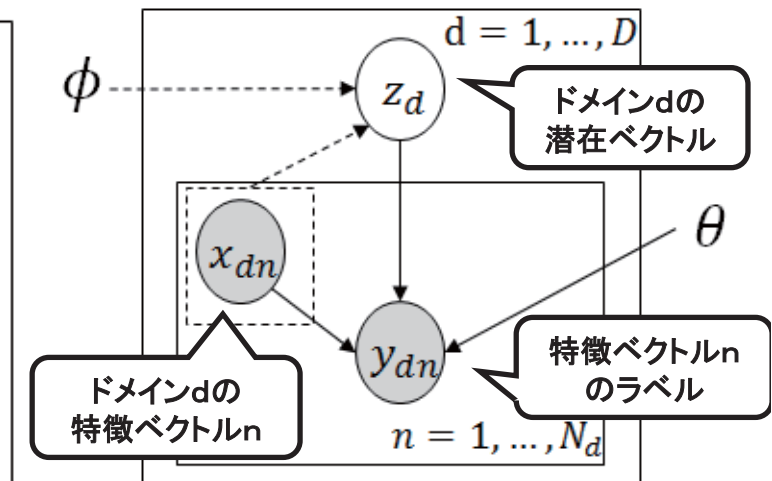
ドメインを表す潜在ベクトルを導入し、テスト時に得られる目標ドメインの特徴ベクトル集合からそれを推論することで、目標ドメインに適したモデルを推論

結果

4種類の実データセットを用いて実験を行い、既存法を上回る性能を示した



潜在ドメインベクトルを介して、特徴ベクトル集合から教師ありモデル(分類器など)へのマッピングを学習



提案法のグラフィカルモデル

→ 提案モデル
 - - - - - 推論モデル

敵対強化学習におけるゲームの複雑性の定量評価

大嶋 真理絵 (NEC-産総研AI連携研) 中田 亨 (産総研・NEC-産総研AI連携研)

目的

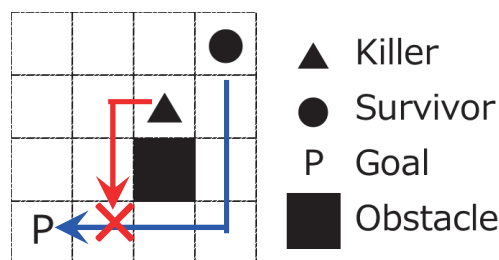
ゲームのルールの情報複雑性を測る

手法

ルール追加による行動選択エントロピーの均衡の変化を測る

結果

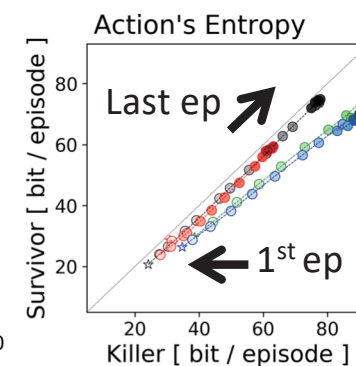
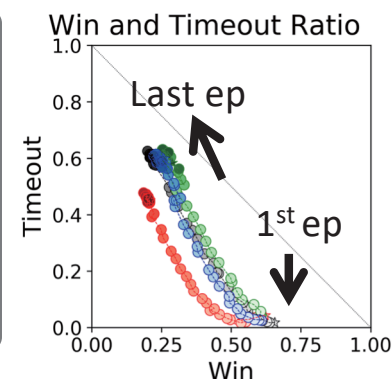
ルールによるゲームの複雑性の変化は単調増加ではない



基本のゲーム (Plain)

追加するルール

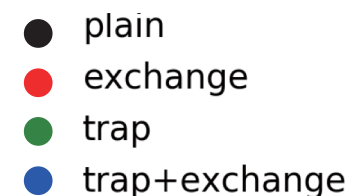
Exchange :
特定の場所で位置が入れ替わる
Trap : 罠を仕掛けられる
Exchange + Trap :
上記の2ルールを同時に追加



ルールを追加するとゲームの複雑性はどうなるか？

結果

- 学習が進むと、行動選択エントロピーは増加する傾向
- ルールによるゲームの複雑性の変化は単調増加ではない
- エントロピーの45度線からの距離はプレイヤー間の方策の複雑性の非対称性を表す



教師付解釈学習

秋田 大空 (京大) 鹿島久嗣 (京大・理研AIP)

目的 深層学習によるモデルの解釈性(interpretability)を向上させたい

手法 ラベルの根拠となる部分も入力として学習，勾配を用いた正則化

結果 精度の向上を人工データ及び実データで確認

- 深層学習によって構成されたモデルは**解釈性が低い**ことが多い
 - 現在は精度の高いモデルに対して後付で説明を与える
“post-hock interpretability” が主流
 - 最初から**精度と解釈性を高めるように学習**する方法はないか？
- データのラベルだけでなく**ラベルの根拠(エビデンス)も入力として使用**
 - エビデンス領域の**予測に対する勾配が大きくなるように学習**
 - **勾配による正則化**を加えたLoss関数や**非エビデンス領域のDropout**を用いて精度及び解釈性の向上をはかる

トラッキングズレを含むデータのマイニング

糊澤 信（AGC株式会社）

目的

アノテーションの紐付が信頼できないデータセットでマイニングをしたい。

手法

試行錯誤的に弁別性が高くなるように近傍でアノテーションを付け替える。

結果

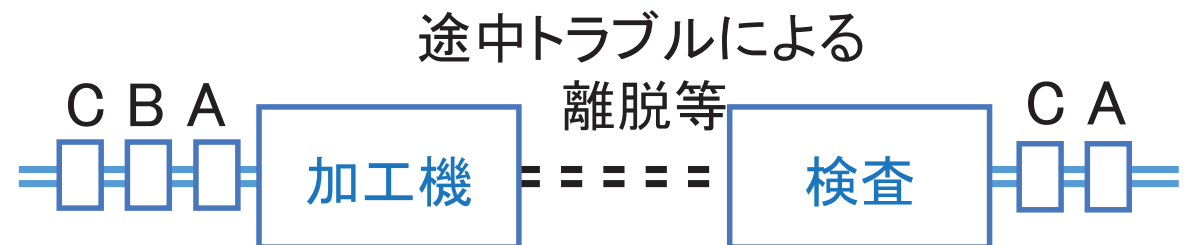
モデルデータを用いた実験で、実現の可能性と有効な範囲を検証した。

【仮定1】トラッキングのズレは
近傍の範囲でしか起こらない

【仮定2】不良品は通常の条件
から外れたときに発生する



OK、NGの弁別性が高くなる
アノテーションの組合せを探索



正しいデータセット

ID	加工データ	検査データ
A	0.29	OK
B	0.20	離脱
C	1.53	NG
D	0.73	OK

得られるデータセット

ID	加工データ	検査データ
A	0.29	OK
B	0.20	NG
C	1.53	OK
D	0.73	OK

トラッキングズレ

正しいアノテーションに復元したい

クラウドソーシングを利用した 階層マルチラベル分類の品質管理

竹岡邦紘, 小山田昌史 (NEC)

目的 回答統合の高精度化

手法 ラベルの階層構造を反映したアノテータのスキルに基づく確率的生成モデル

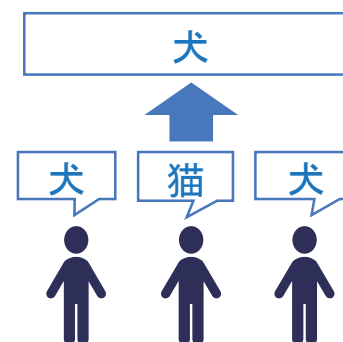
結果 精度の高い回答統合と、理解可能なアノテータのスキルが得られる

着眼点

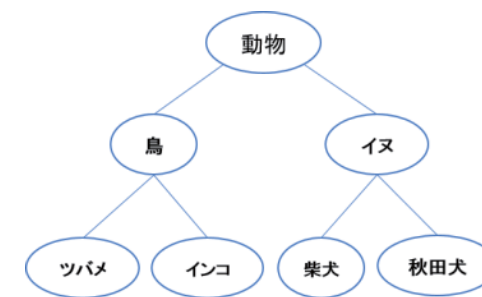
- 与えられるラベルの階層構造を活用
- アノテータはラベルに対して得手不得手がある

モデルのエッセンス

- 「柴犬」のラベルに対して詳しい(スキルの高い)アノテータは、「柴犬」に類似した「秋田犬」のラベルについても詳しい
- スキルとタスクの相性に応じて、アノテーション結果の信頼度が変化



回答統合



ラベルの階層構造

過剰エントロピー生成最小化としての 拡張Contrastive Divergence法によるRNNの学習

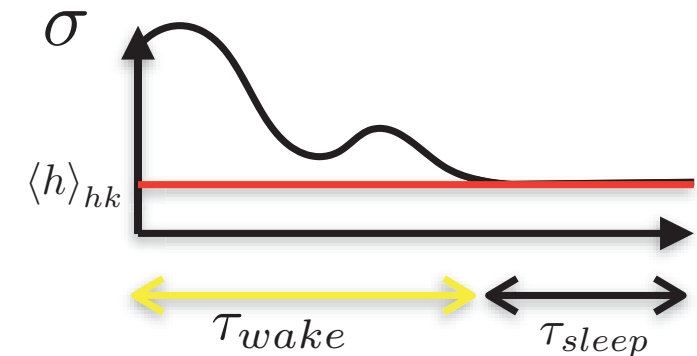
山本峻平 (東大) 金子邦彦 (東大)

目的 RNNの新しい学習則を提案

手法 熱力学的な量を擬似的に最小化

結果 BPTTを用いないHebb型学習則により実装可

- Contrastive Divergence法は非平衡統計力学の枠組で解釈可能
 - Contrastive Divergenceは緩和のエントロピー生成に等しい
 - エントロピー生成のうち発熱だけを用いた擬似的SGDがCD法
- 系の定常状態を非平衡, 入力を時系列データにして確率的RNNへ拡張
 - 入力による仕事量と過剰エントロピー生成の和を0にする学習
 - エントロピー生成からhouse-keeping項を差し引いたものが過剰エントロピー生成
 - 「睡眠」過程を用いることで過剰エントロピー生成も発熱だけで最小化可能
 - BPTTを用いない局所的なHebb則だけで構成できる学習アルゴリズムとなる



Direct Feedback Alignmentの解析

市川航平、金子邦彦（東京大学総合文化研究科）

目的

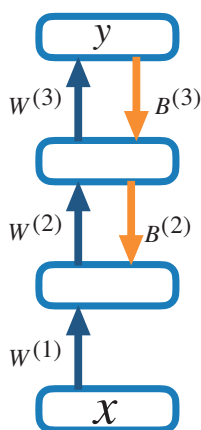
biologically plausible back propagationの一つとして提案されているDirect Feedback Alignmentの理解を深める。

手法

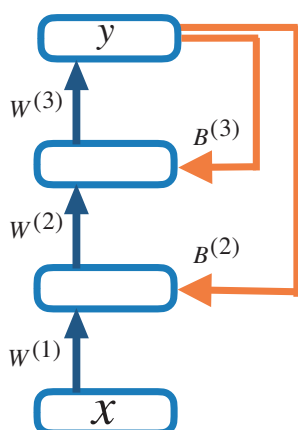
- Linear classifier probeによって隠れ層の寄与度を測定し、DFAを説明できると提案された手法であるLAFSとの比較を行った。
- 隠れ層への入力分散に着目し、DFAと類似の手法であるFeedback Alignmentとの比較を行った。

結果

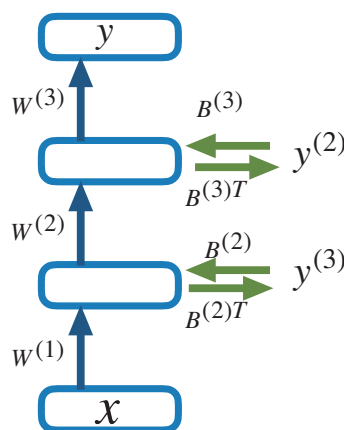
- DFAは実質的に各層毎に個別の学習をしていることに対応するのではないかと指摘された(ICLR2017)が、そうではなくDFAはNNの多層性を利用できていることが示唆されるような結果を得た。
- DFAはFAを拡張したものであり、より計算を簡略化した手法であるがその性能にはほとんど差がない。上の結果を踏まえ、その理由としてFAは隠れ層への入力分散が小さくなる（=NN全体が線形システムに近づく）ような制約が暗に働き、結果的にNNの表現力が小さくなっているのに対してDFAではその問題が回避されているからであろうという仮説を得た。



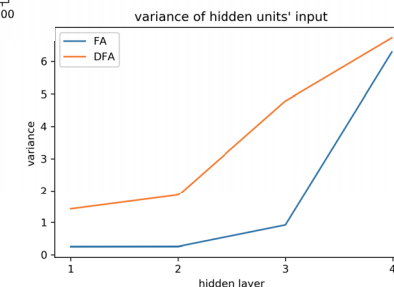
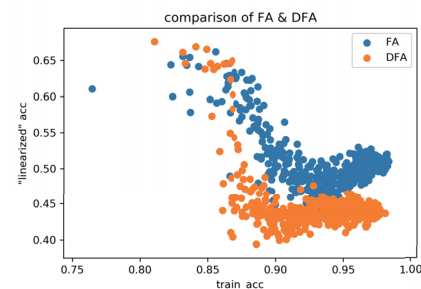
FA



DFA



LAFS



On the Minimal Supervision for Training Any Binary Classifier from Only Unlabeled Data

Nan Lu (UTokyo) Gang Niu (RIKEN-AIP)

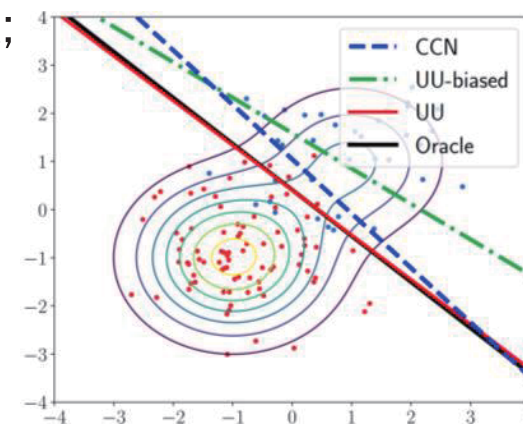
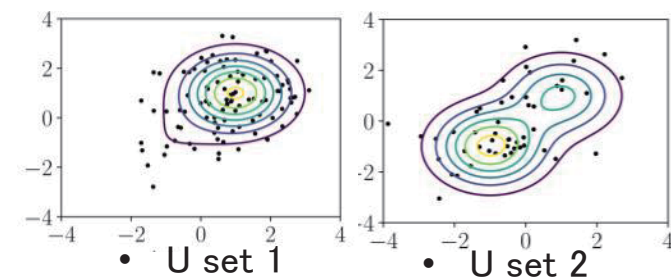
Aditya K. Menon (ANU) Masashi Sugiyama (RIKEN-AIP • UTokyo)

目的 Learn arbitrary binary classifier (linear to deep) from only unlabeled (U) data.

手法 Propose the first empirical risk minimization based learning method.

結果 Three class priors are all we need to train any binary classifier from only U data.

- In big data era, labeling cost is extremely expensive, can we learn from only U data?
- Unsupervised classification (clustering)
 - Cluster assumption $\rightarrow$ rarely holds in practice.
- Proposed two-step approach (UU method)
 - Setp1: estimate the risk from two sets of U data;
 $\rightarrow$ leads to unbiased risk estimator
 - Step2: empirical risk minimization.
 - Without additional distributional assumption.
- Learn good classifier even without any label!



- Classification results, UU in red line is the proposed method, we can see UU outperforms other methods.

Co-teaching: Robust Training Deep Neural Networks with Extremely Noisy Labels

Bo Han Quanming Yao Xingrui Yu Gang Niu,
Miao Xu Weihua Hu Ivor Tsang Masashi Sugiyama

目的

手法

結果

Masking: A New Perspective of Noisy Supervision

Bo Han Jiangchao Yao Gang Niu Mingyuan Zhou,
Ivor Tsang Ya Zhang Masashi Sugiyama

目的

手法

結果

ベイズ最適化を用いた多次元ヒストグラムのビン数最適化

武藤健介（東大） 坂本浩隆（東大） 永田賢二（産総研）
有馬孝尚（東大） 岡田真人（東大）

目的

多次元ヒストグラムのビン数を，自動的かつ高速に推定する手法を構築すること．

手法

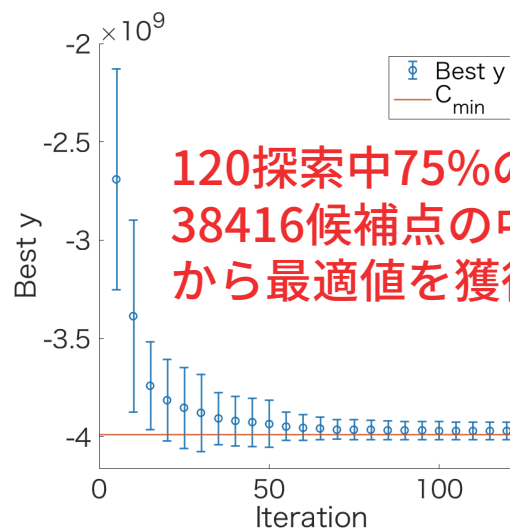
コスト関数の計算による，多次元ヒストグラムのビン数推定と，ベイズ最適化（BO）を用いた計算量の削減．

結果

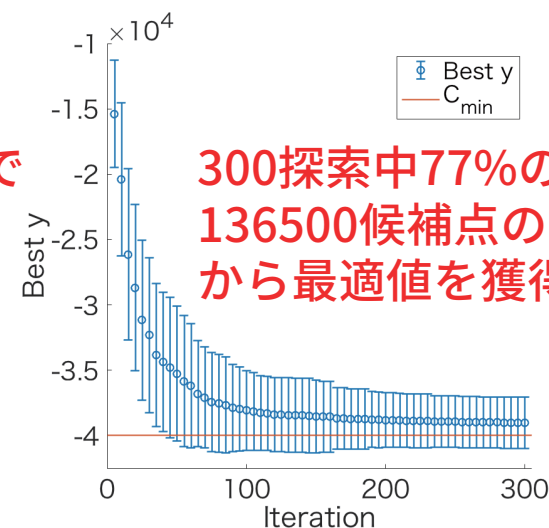
4D人工データ・4D実データに対して提案手法を適用したところ，BOを用いる事によって探索回数を数百倍削減できた．

- 4次元コスト関数 $C(N)$ に対するBOの適用結果（探索回数と獲得した $C(N)$ の最小値）
- 人工データ・実データに対して，同じ実験を100回行いエラーバーを付けた．
- 赤線は $C(N)$ の最適値を表す．

BOの結果（人工データ）



BOの結果（実データ）



単語ベクトルを利用した文書の教師なし行列表現

福井 一輝 (京都大学・理研AIP) 下平 英寿 (京都大学・理研AIP)

目的

訓練済みの単語ベクトルを用いて文書間の距離を測る

手法

語順をランダムに入れ替えた文と元の文との違いを捉える

結果

文書内の単語の位置関係の情報が文書間の距離に反映される

文書内の単語ベクトルを並べた行列

$$S = (w_1, w_2, \dots, w_n)$$

ランダムに
語順を変える

$$\bar{S}^{(1)}, \bar{S}^{(2)}, \dots, \bar{S}^{(k)}$$

S を変換

$S, \bar{S}$ の違いを
捉えるような
線形変換を学習

$$\tilde{S} := A_L S A_R$$

変換された行列どうしの
距離には単語の
位置関係が反映される

テクスチャ画像識別問題に対する フーリエ変換を用いたデータ拡張の検討

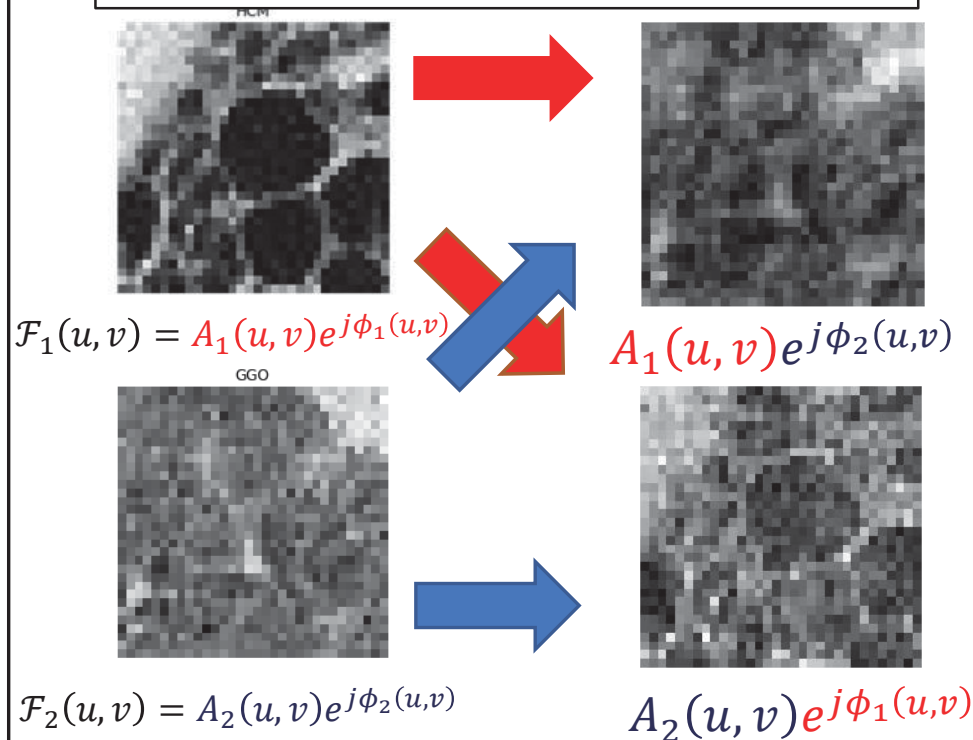
新田 大悟 (電通大) 庄野 逸 (電通大)

目的 テクスチャ画像に対するデータ拡張

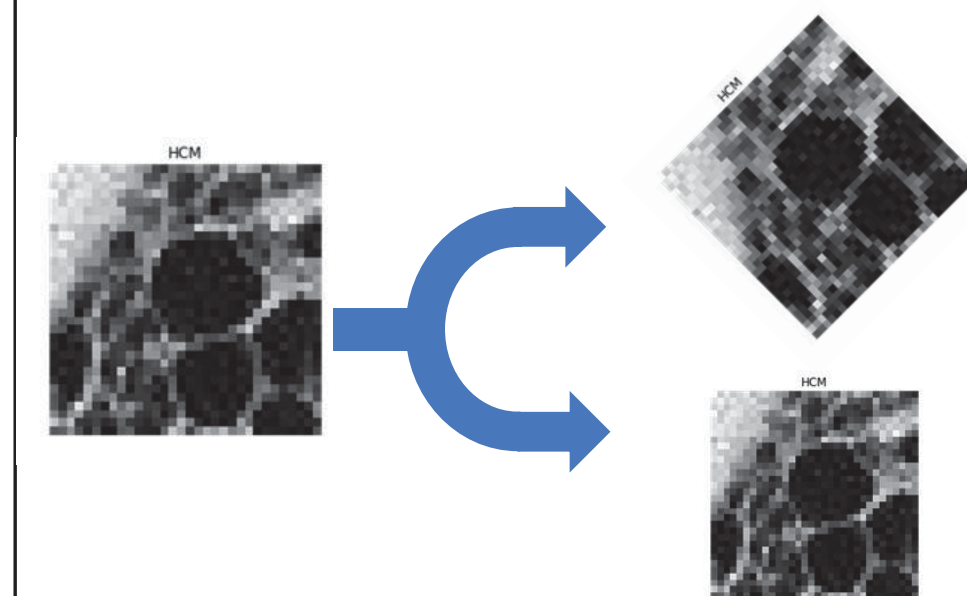
手法 フーリエ変換を用いて画像を分解し, その振幅を別クラス画像の位相と合成

結果 線形変換によるデータ拡張と比較して識別精度が向上した

フーリエ変換によるデータ拡張 (識別率:88.5%)



線形変換(回転)によるデータ拡張 (識別率:78.3%)



$A(u, v)$: 周波数振幅特性
 $\phi(u, v)$: 周波數位相特性

強化学習における未知環境での最適方策の 効率的推論に向けた検討

榎本 昌文 岡留 剛 (関西学院大学)

目的

MDPの最適状態価値関数を, 与えられたMDP群を用いて推論.

手法

MDPをLMDPに埋め込み, 埋め込み先が一番近いMDPで代用.

結果

MDPの最適状態価値関数を別のMDPで代用可能.

目的

given MDPs

- β
最適状態価値関数 v_β^*
遷移確率 P_β が既知

transfer

target MDP

- α
遷移確率 P_α も既知

効率良く v_α^* を計算したい.

手法

LMDP space

given MDPs

■ β

target MDP

■ α

埋め込み先が近いMDP β の v_β^* で代用

結果

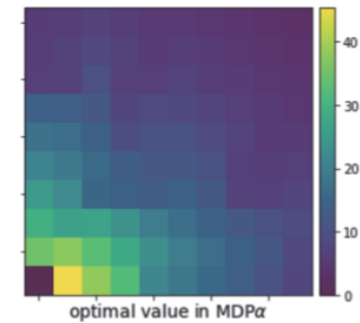
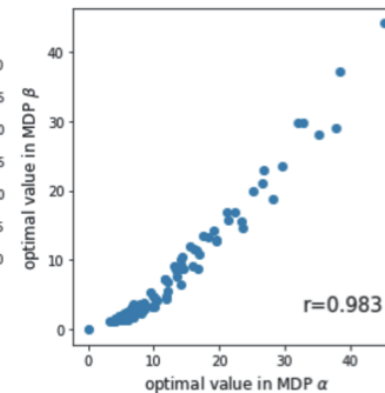
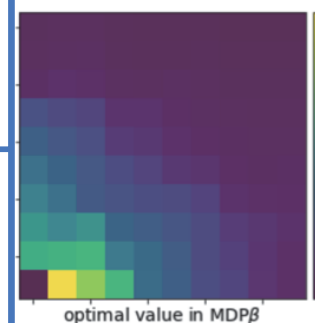
LMDP space

given MDPs

■ β

target MDP

■ α



v_α^* と v_β^* に強い相関があることを確認
定数倍すれば代用可能

Total Stochastic Gradient Algorithms with Application to Model-based Reinforcement Learning

Paavo Parmas (Okinawa Institute of Science and Technology)

目的

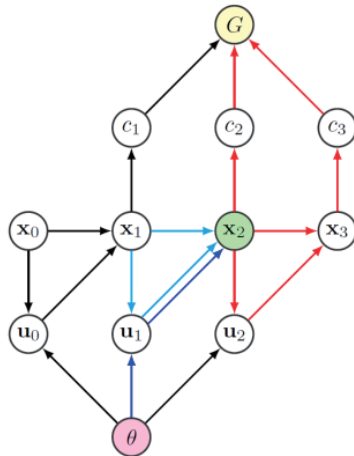
Flexible model-based reinforcement learning by differentiating predictions

手法

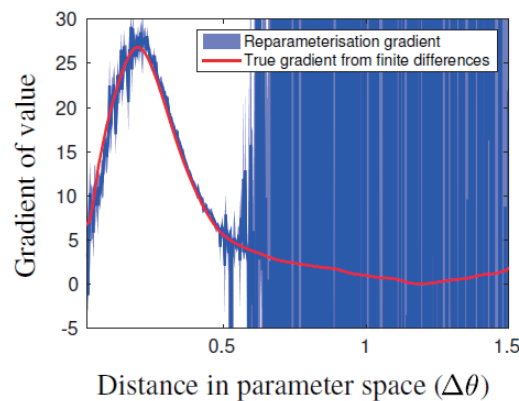
Created new theoretical framework and gradient estimation techniques

結果

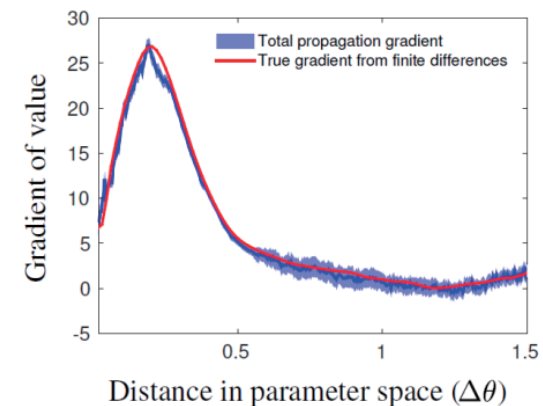
Gradient accuracy is sometimes 10^6 times better, and learning works



New framework



Old method



New method

[1] Parmas, P., Rasmussen, C. E., Peters, J., & Doya, K. (2018, July). PIPPS: Flexible Model-Based Policy Search Robust to the Curse of Chaos. In *International Conference on Machine Learning (ICML)*.

[2] Parmas, P. (2018, Dec). Total stochastic gradient algorithms and applications in reinforcement learning. In *Advances in Neural Information Processing Systems (NIPS)*.

テンソルネットワークを用いた教師あり学習

新谷祐矢（神戸大） 笹川佳則（神戸大）

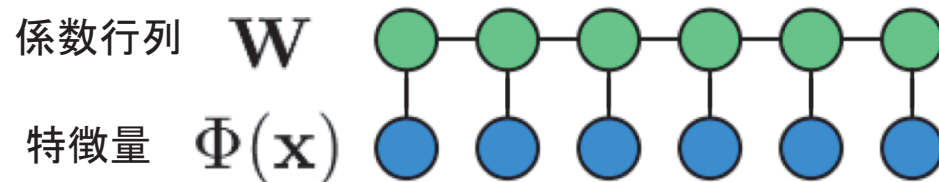
目的 量子多体系の計算手法を用いた教師あり学習法の構築

手法 決定関数に含まれる高次元テンソルのテンソルネットワークによる近似

結果 構築したモデルは既存手法と同等またはそれ以上の汎化性能をもつ

• 決定関数

$$f(\mathbf{x}) = \mathbf{W} \cdot \Phi(\mathbf{x})$$

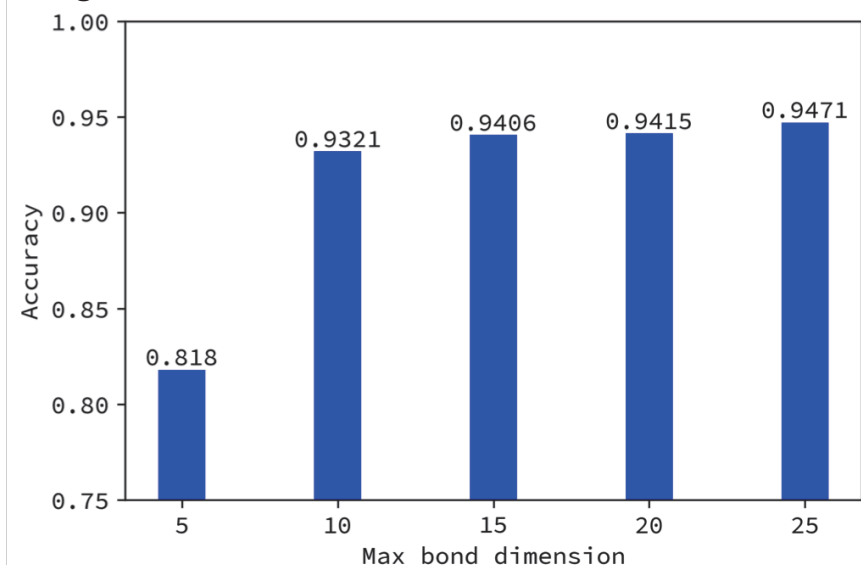


• 学習方法

隣り合うふたつのテンソルを
勾配降下法を用いて同時に更新する

• 分類問題への適用

Fig. 14x14 手書き数字の分類精度



Max bond dimension : 近似精度を調節するhyper parameter

ガウス過程回帰を用いた適応的物性スペクトル測定

大塚 琢馬 若林 勇希 谷保 芳孝 山本 秀樹 澤田 宏

目的

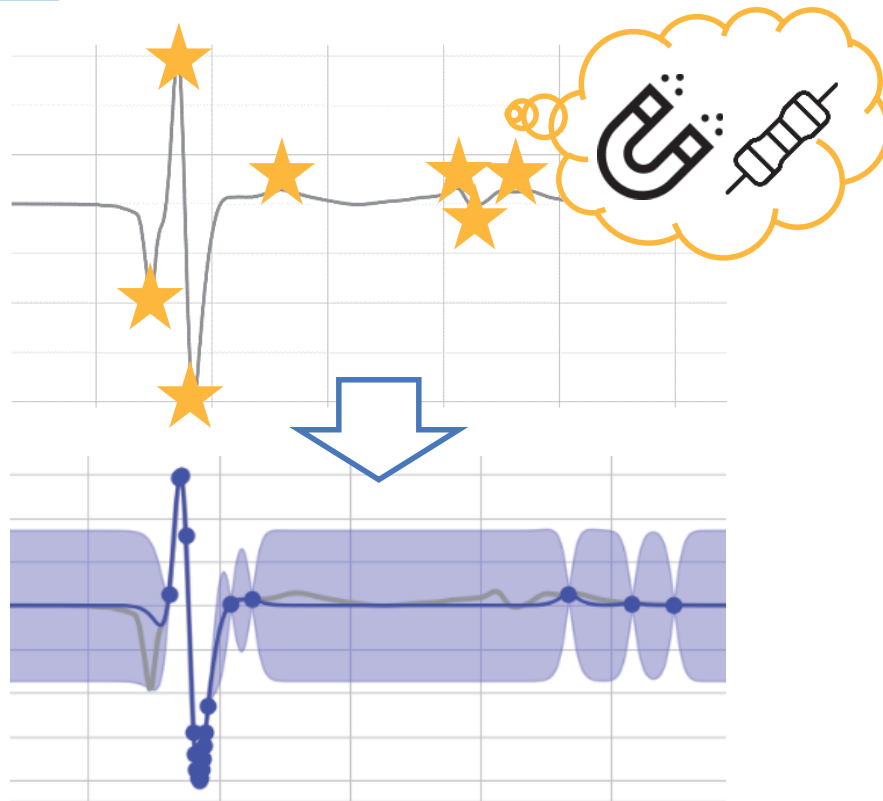
限られた観測点から物性スペクトルをガウス過程を用いて復元

手法

スペクトルの勾配や振幅の予測に基づき観測点を逐次選択

結果

特徴の異なる種々のスペクトルで勾配や振幅情報の特性を確認

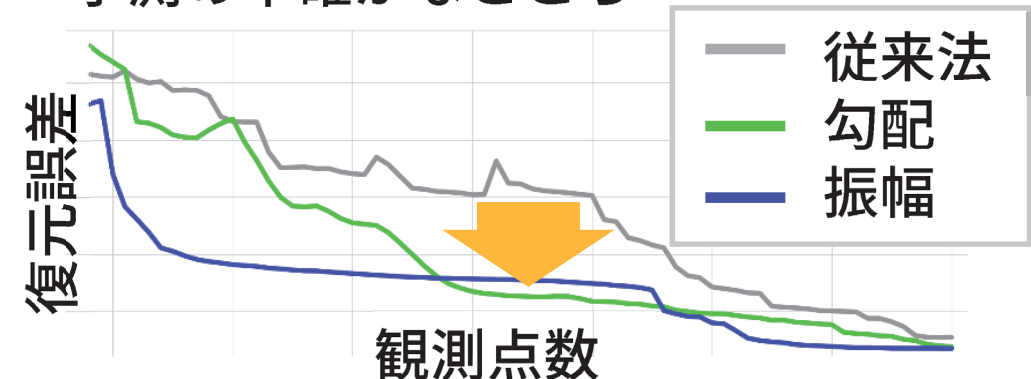


物性スペクトル

- 物質の光・電気・磁氣的性質に対応
- 特にピーク★に現れる
- 網羅的な観測を避けて効率化したい

ピークを効果的に選ぶには?

- 勾配，振幅の大きいところ
- 予測の不確かなところ



熱浴法の幾何学的描像

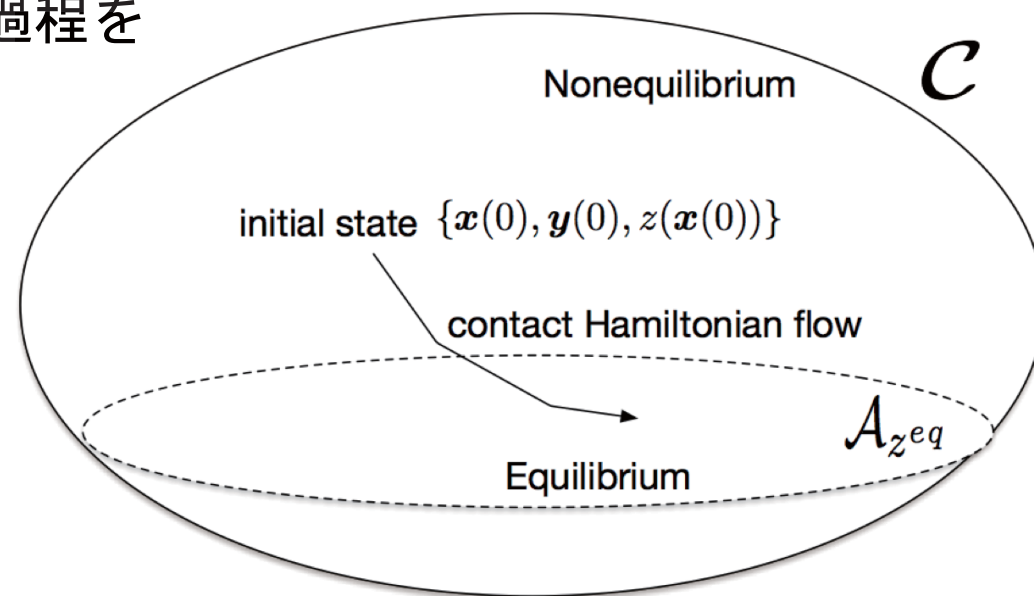
日野 英逸 (統数研) 後藤 振一郎 (京大)

目的 熱力学的緩和過程と平衡状態の統一的な幾何学的記述

手法 力学系理論：ハミルトンベクトル場とMaster方程式

結果 理想的な熱浴法が接触/情報幾何的な緩和過程と対応することを見出した

- シンプレクティック幾何学がハミルトン力学系の記述に用いられるように、熱力学的マクロ変数の力学系は接触幾何で定式化すると見通しが良い
- 非平衡状態から平衡状態への緩和過程を表す力学系は少し前に導かれた
- 具体的にその力学系を実現するMaster方程式とポテンシャルの組として、ある理想的な熱浴法が合致することを見出した
- 展望：MCMCの幾何学的解析と良いアルゴリズムの導出



カーネル密度関数推定を用いたパラメータ密度の継承による CNNの学習効率化

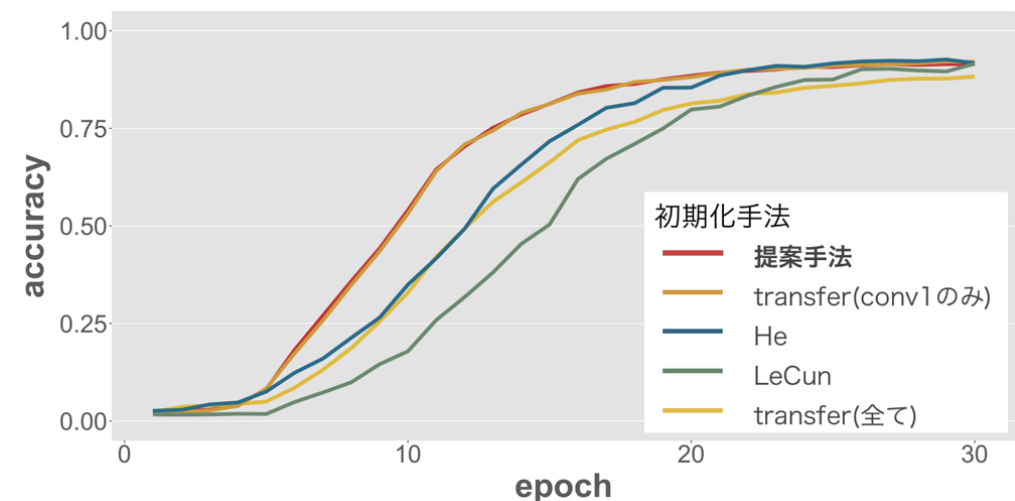
堀内 圭佑 (筑波大) 亀山 啓輔 (筑波大)

目的 畳み込みニューラルネットワークの学習効率化

手法 カーネル密度推定を用いた、訓練済みネットワークのフィルタベクトルの確率密度の継承によるCNNの初期化

結果 既存の初期化手法に対する学習効率の向上

- 本研究では、CNNの学習の効率化を目的として、KDEを用いたネットワークの初期化手法を提案する。提案手法では、学習済みネットワークのフィルタの確率密度を継承することで、新たなネットワークの学習を効率化し、異なる構造のネットワークに対しても適用できる。実験では、既存手法に対して改善が見られた。



k-近傍交叉カーネルを用いたNadaraya-Watson回帰の高速近似計算法

伊東 利雄 濱田 直希 大堀 耕太郎 樋口 博之(富士通研究所)

目的

k-近傍交叉カーネル回帰法の推定計算時間の短縮化

手法

推定計算の近似計算法を提案

結果

推定精度を落とすことなく高速に計算することが可能

- 従来の推定計算方法の特徴
 - サンプルセットのすべてのサンプル点から作成されたカーネル関数を使用
- 新近似計算法の特徴
 - 推定を行いたい点の近傍内にある, サンプル点から作られたカーネル関数のみを使用
- 性能評価結果(Power Plantのデータセット)
 - 近似計算法により, 従来の推定計算方法に対し, 推定精度を落とすことなく, 8.7倍高速に計算が可能になり, 計算時間が短縮される

MR画像再構成における複数の画像事前分布を組み合わせた 事後平均近似解

久保田 菜々子(早大) 原田 賢(早大)
藤本 晃司(京大) 岡田 知久(京大) 井上 真郷 (早大)

目的 事後平均推定による二乗誤差の縮小を目指す。

手法 混合分布族を指数分布族に近似することで事後平均近似解を求める。

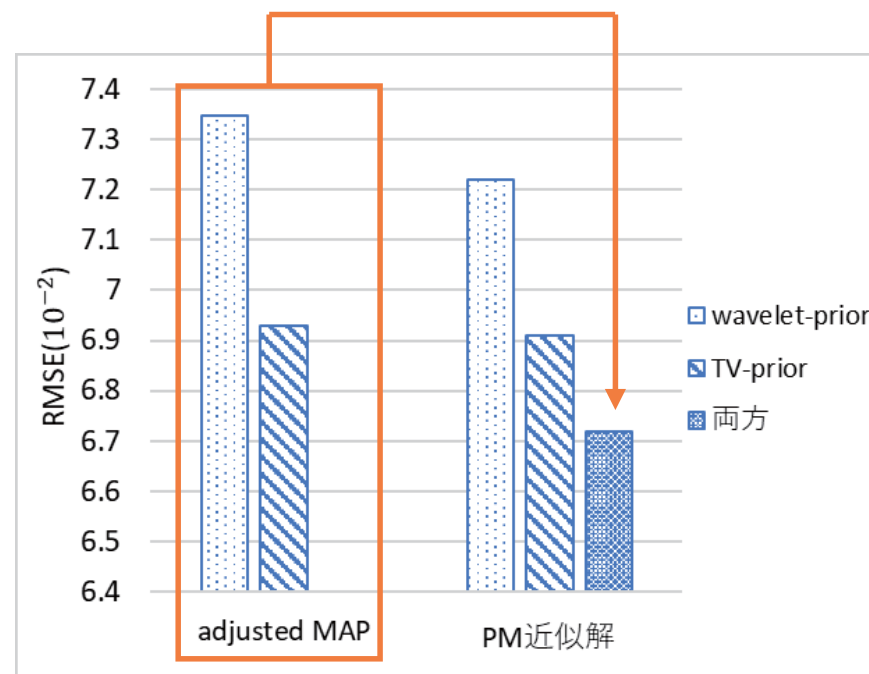
結果 計算量のオーダーを変化させずに精度を向上させた。

単独の画像事前分布により再構成したMAP解を複数組み合わせることで事後平均解に近似
MAP解(指数分布族)

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \underbrace{\frac{1}{2\sigma^2} \sum_{c=1}^C \|\mathbf{y}_c - \mathbf{F}_u \mathbf{S}_c \mathbf{x}\|_2^2}_{\text{観測モデル}} + \underbrace{\sum_{k=1}^K w_k f_k(\mathbf{x})}_{\text{複数の画像事前分布}} \right\}$$

事後平均近似解(混合分布族)

$$\begin{aligned} \tilde{\mathbf{x}} &\equiv \langle \mathbf{x} \rangle_{p(\mathbf{x}|\mathbf{y};\mathbf{w})} \\ &= \sum_{k=1}^K \hat{w}_k \langle \mathbf{x} \rangle_{p(\mathbf{x}|\mathbf{y},k)} \cong \sum_{k=1}^K \hat{w}_k \hat{\mathbf{x}} \end{aligned}$$



混合ロジスティック回帰の局所変分近似による推定と変分自由エネルギーの考察

中村 文士 小西 亮介(ゼネリックソリューション) 清木 康(慶應)

目的 混合ロジスティック回帰モデルを精度良く、効率的に推定したい

手法 局所変分近似により近似的なベイズ事後分布を求める

結果 従来法よりも精度良く、計算量の少ない推定が可能になった

混合ロジスティック回帰モデル:

$$p(y|x, \{a_k, b_k\}_{k=1}^K) = \sum_{k=1}^K a_k r(x, b_k)^y (1 - r(x, b_k))^{1-y}$$

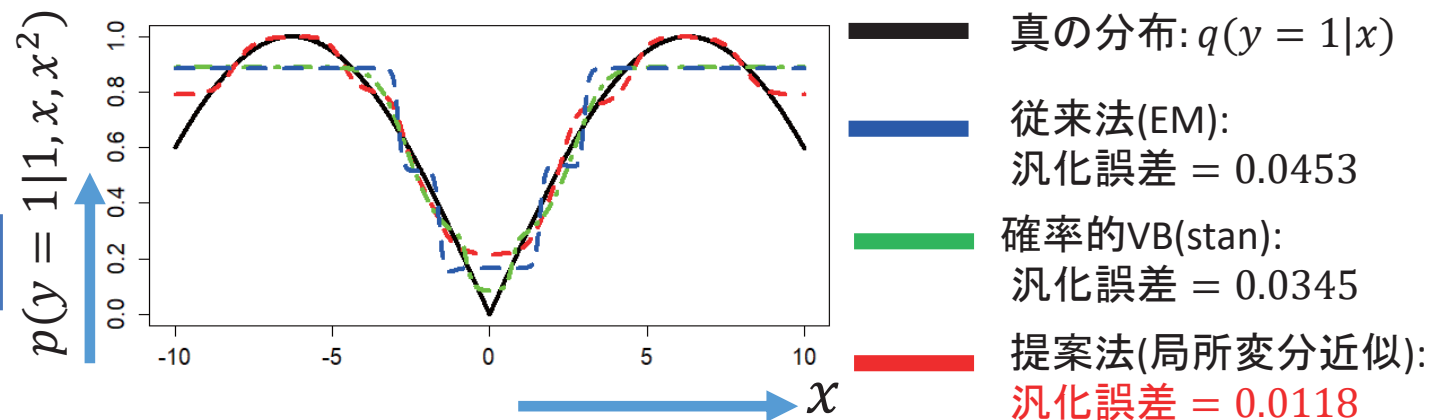
$r(\cdot)$: sigmoid function

共役な事前分布がないので効率的に計算するのが難しい

→局所変分近似が適用できそう、アルゴリズムを導出した

変分自由エネルギーの結果はポスターで!!!

予測分布の比較(赤:提案法)



独立低ランクテンソル分析： 非負値性・低ランク性・独立性に基づくブラインド音源分離の統一理論

吉井 和佳 (京大・理研) 佐藤 寛之 (京大) 坂東 宜昭 (京大・産総研)
中村 栄太 河原 達也 (京大)

目的

音響信号 (時間-周波数-空間軸のテンソル) に対するブラインド音源分離

手法

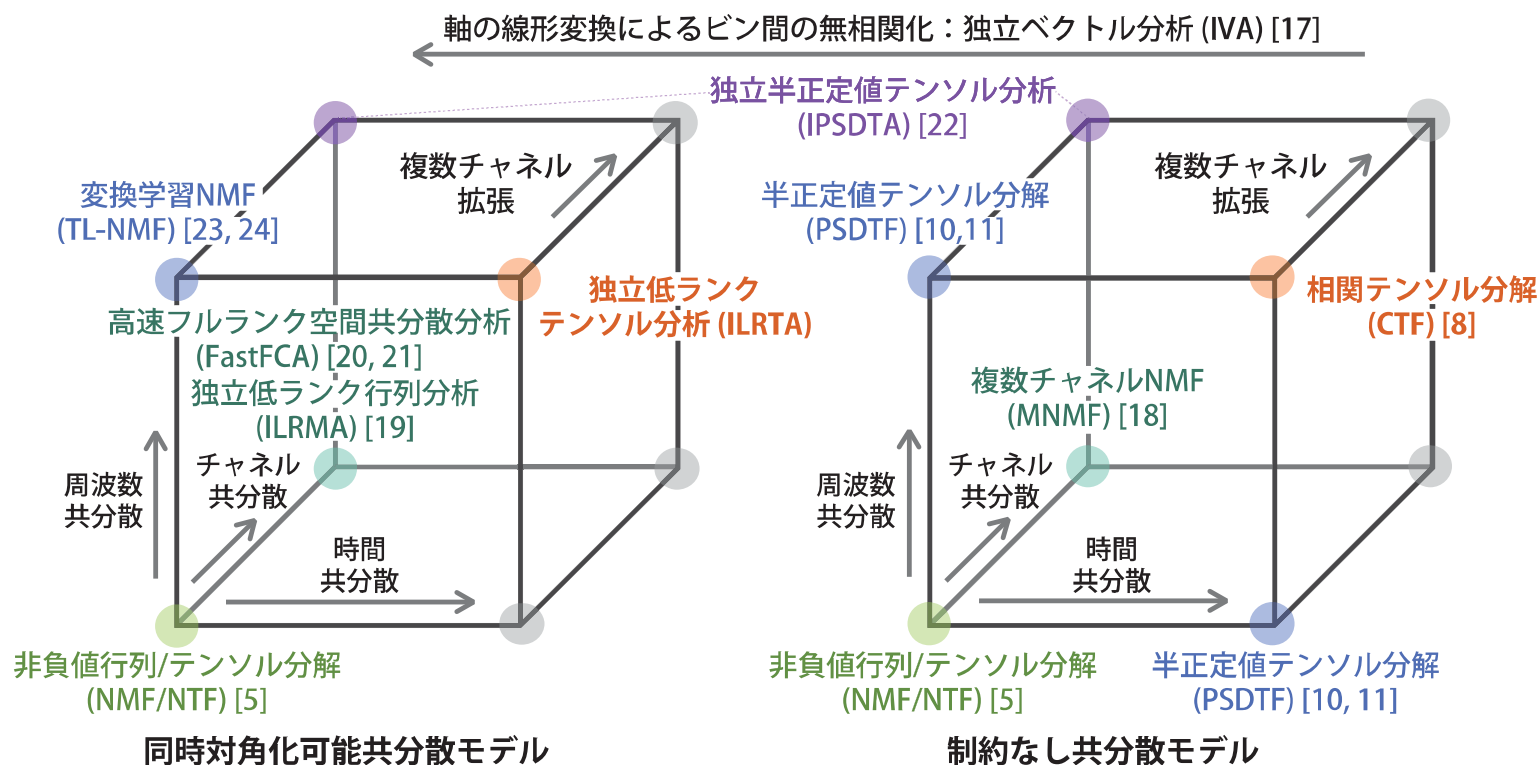
テンソル中の全要素間の共分散行列に対する低ランク分解 (CTF)

→ ILRTA: 各軸の線形変換 (IVA) + 変換後テンソルの低ランク分解 (NTF)
近似

結果

主要な音源分離法を CTF or ILRTA の特殊形として統一的に説明

高速



f 平均によるディリクレ過程平均法の一般化と 影響関数の解析

小林 真佐大（豊橋技科大） 渡辺 一帆（豊橋技科大）

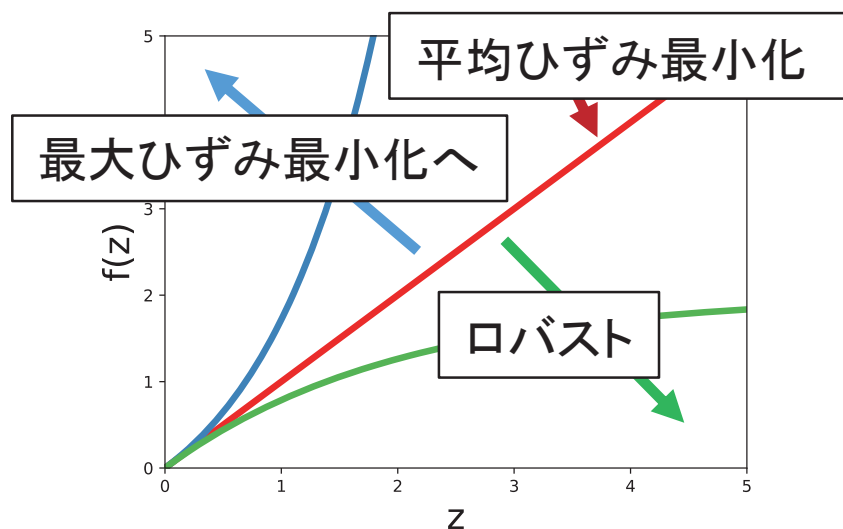
目的 最大ひずみの最小化および外れ値に対する頑健性の獲得

手法 f 平均により目的関数を一般化

結果 一般化アルゴリズムの導出および影響関数の導出

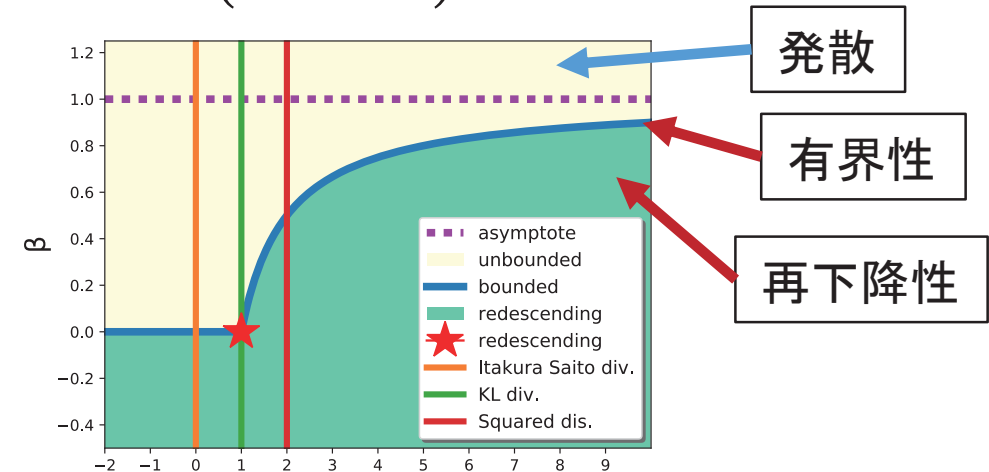
1. 目的関数の一般化

$$f^{-1} \left(\frac{1}{n+K} \left\{ \sum_{i=1}^n f(d_{\phi}(x_i, \theta_{c(i)})) + f(\lambda)K \right\} \right)$$



2. 影響関数の導出

$$\text{IF}(x^*) \propto f'(d_{\phi}(x^*, \theta)) (-\nabla \nabla \phi(\theta)(x^* - \theta))$$



β ダイバージェンス(図中の α)に対する
影響関数の挙動

非負値行列分解における変分近似精度の理論解析

林 直輝（株式会社NTTデータ数理システム）

目的 非負値行列分解(NMF)における変分Bayes法(VB)の近似精度の理論を構築

手法 VBNMFと同じモデルでNMFの真と学習器のKL情報量を数学的に解析した

結果 NMFにおけるVBとBayes推測についてその精度の差異の挙動を導出した

• 推測精度の理論値は……

——	事前分布	
手法	零点なし	ガンマ分布
VB推測	手法なし	解明済み By Kohjima
Bayes推測	解明済み By Hayashi	<u>????</u>

- M. Kohjima and et al. , *ICANN*, 2017
- N. Hayashi and et al. , *Neucom.* , 2017
- N. Hayashi and et al. , *IEEE SSCI*, 2017

- 主定理 1 :
事前分布がガンマ分布（ $\exists$ 零点）
のときの推測精度の理論上界を導出
- 主定理 2 :
変分近似精度=VBとBayesの推測精度の差
理論下界を導出: $\bar{F}_n - F_n \geq \tilde{\lambda} \log n + O_p(1)$

$$\tilde{\lambda} = \begin{cases} \text{????} & \text{phase 1} \\ \text{????} & \text{phase 2} \end{cases}$$

- 詳細はポスターT-08へお越しください！

マルコフネットワーク構造変化検出における スパースKLIEPモデルのセーフスクリーニング

佐藤 浩基（岐阜大） 志賀 元紀（岐阜大・JST・理研）
山田 誠（京大・JST・理研）

目的 マルコフネットワーク構造変化検出モデルKLIEP^[1]の効率的な計算

手法 KLIEPモデルに対する**セーフスクリーニング法**

結果 KLIEPモデルの**安全かつ効率的な特徴選択**

状態の変化に伴うネットワーク構造変化箇所の特定

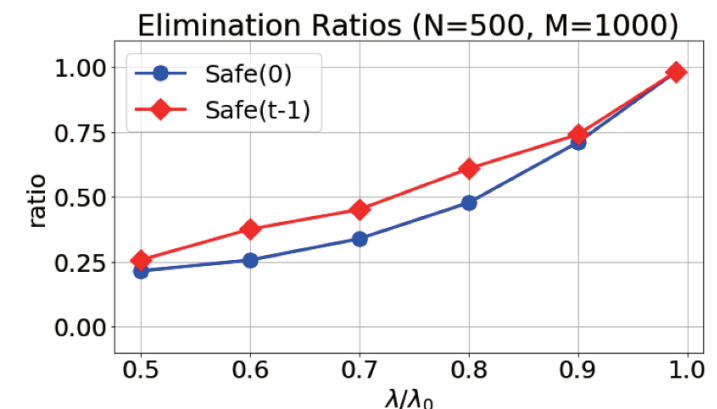
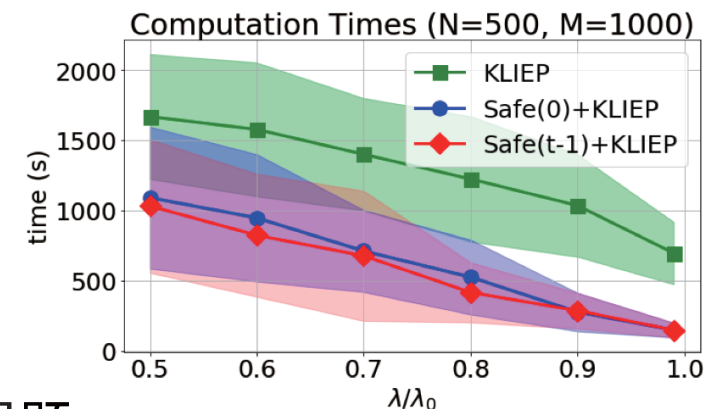
➤ 癌細胞と正常細胞の遺伝子相互作用の差異

膨大な推定パラメータの量（ネットワークのエッジ数）が問題

KLIEPモデルのセーフスクリーニング法を開発

➤ 特徴量毎に**並列処理**可能, **安全な**特徴選択

➤ 理論的に導出, 実験的に有効性を確認



[1] Song Liu, Kenji Fukumizu, and Taiji Suzuki. *Behaviormetrika* 44.1 (2017): 265-286.

近傍法と形式概念解析を用いた階層的構造の学習

米田 友花(産研, 阪大) 杉山 磨人(NII) 鷺尾 隆(産研)

目的

多変量データから重複するクラスターの階層構造を発見する

手法

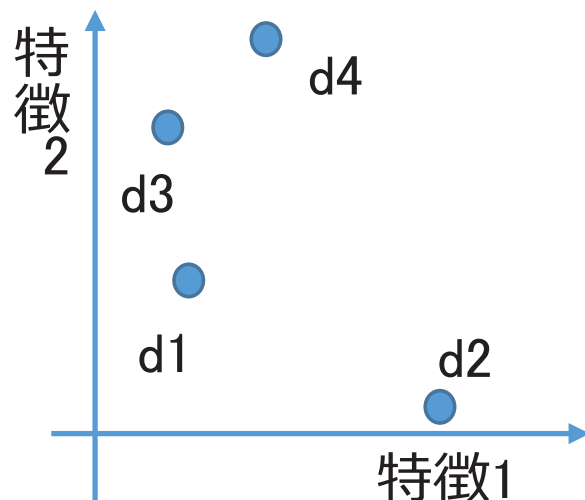
近傍法と形式概念解析を組み合わせた

結果

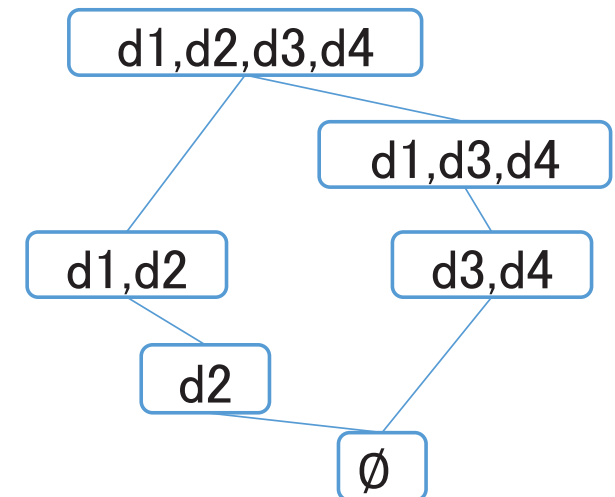
典型的手法より質の良い階層クラスターが学習できた

①近傍法による二値化: 元の特徴空間での局所的なデータ点の近さに基づく

②形式概念解析: 階層的構造の構築, 飽和アイテム集合を利用, 2値データを扱う



トランザクション データ	
X1	d1, d3
X2	d1, d2
X3	d3, d4
X4	d3, d4



Layer Normalizationを用いた単層パーセプトロンの統計力学的解析

高木 志郎 (東大) 吉田 雄紀 (東大) 岡田 真人 (東大)

目的

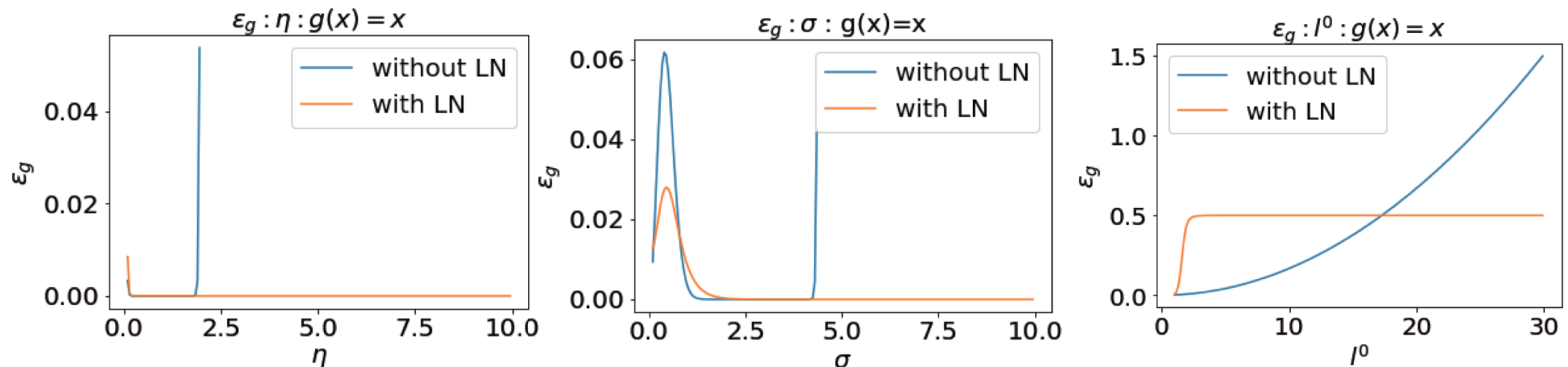
Layer Normalizationがニューラルネットに与える影響を調べる

手法

統計力学的定式化による解析および数値実験

結果

学習係数や重みの初期値，入力分散に対する汎化誤差の頑健性を確認



Layer Normalizationつき単層パーセプトロンの学習ダイナミクスを導出し，パラメータに対する頑健性を確認した

カーネル密度関数の局所変形に基づくトポロジー保存可能な イメージアライメント手法の開発

綿島 正剛 (九工大) 久下 小百合 (九大) 石原 健 (九大)
飯野 雄一 (東大) 吉田 亮 (統計数理研) 徳永 旭将 (九工大)

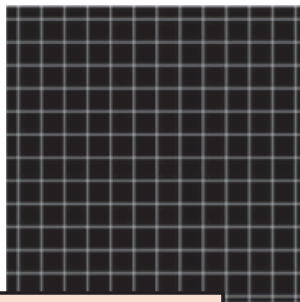
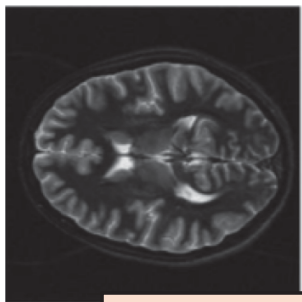
目的 迅速な医療画像診断を支援する汎用的なイメージアライメント手法の開発

手法 カーネル密度推定, EMアルゴリズム, ベイズ最適化

結果 自由変形する管状構造画像に対して高精度の補正を実現

参照画像

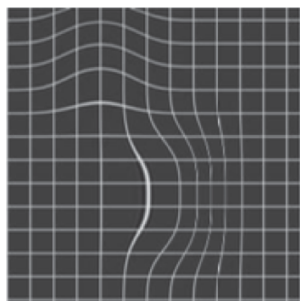
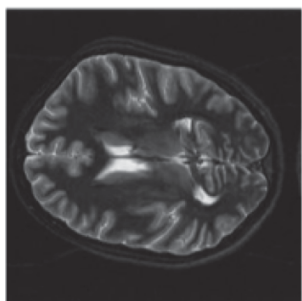
変形場



最適な変形場を推定

実験画像

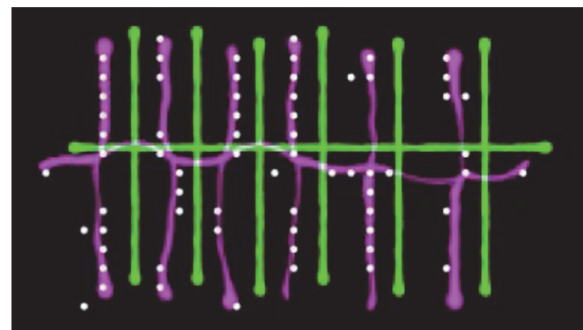
変形場



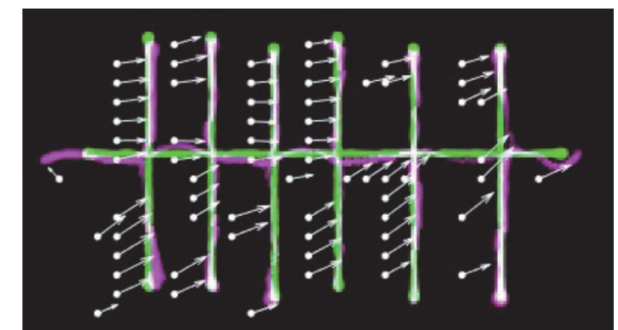
管状構造画像のイメージアライメント

管状構造を持つ, 血管・脊椎画像のアライメントで
画像間の誤差を定量化し, 病変の発見を支援したい

アライメント前の参照/実験画像



アライメント後



■ : 参照画像
■ : 実験画像

○ : 制御点
→ : 変形場 (制御点の移動)

CNNの大規模データセットでのラベルノイズ耐性について

幡谷龍一郎・中山英樹（東大）

目的

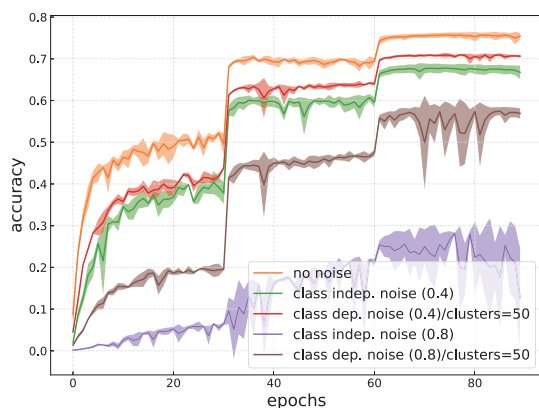
ラベルが全て正しいとは限らない実際の画像データに対してCNNはどこまで頑健か？

手法

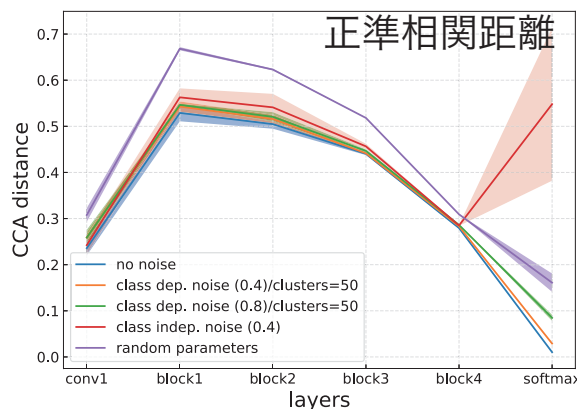
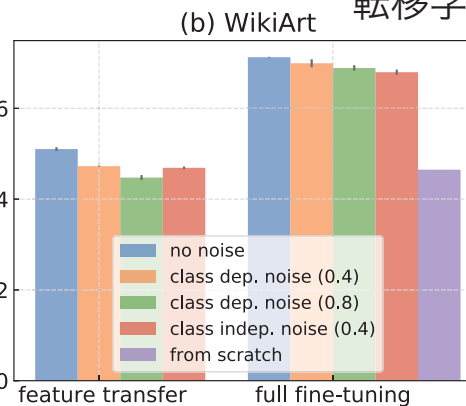
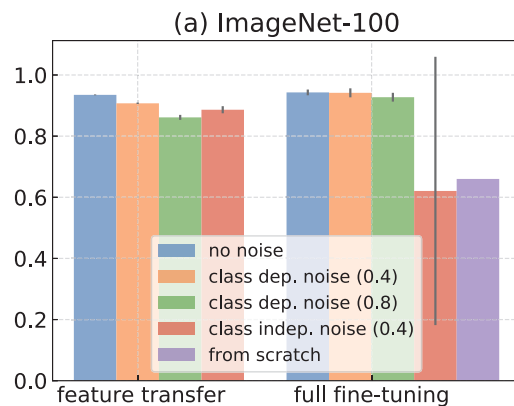
概念間距離を用いてラベルノイズをモデル化し、ノイズつきImageNet-1kを分類

結果

モデル化されたラベルノイズに対してCNNは頑健、ノイズなし時と似たような特徴を学習
転移学習



モデル化されたノイズ(class dep. noise)は一様なノイズ(class indep. noise)よりも精度の低下を引き起こしにくい



- CNNの中間表現間の正準相関距離
- 転移学習後の性能の観点ではモデル化されたノイズはノイズなし時に近い表現を学習する

まとめ

- 一様なノイズは設定として単純すぎる
- CNNは実データのノイズに対してはロバストかも？

ベイズ的最適化を用いた最適集団誘導探索

清水 仁 大塚 琢馬 岩田 具治 澤田 宏 納谷 太 上田 修功 (NTT)

目的

大規模イベントにおいて観客を適切に誘導する方法を見つける

手法

ベイズ的最適化を用いて人流シミュレータ上で効率的に誘導策を探索する
人工データを用いた実験で、全探索よりも約 $10^{23\sim30}$ 倍高速に、かつ、
最短経路よりも優れた誘導策を探索できた

結果

ベイズ的
最適化

① 始点から終点までの
経路を列挙

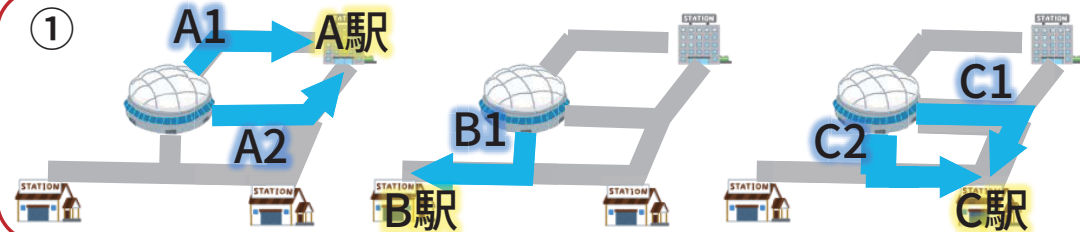
② 各経路に誘導する
人流の配分を設定

③ シミュレーションで
誘導策の効果を評価

④ シミュレーション未実施の
誘導策の効果を予測

⑤ シミュレーション実施した
誘導策の中で最善策を選択

①



②

誘導策の
組合せ爆発

目的地	A駅	B駅	C駅
経路1	0.8	1.0	0.3
経路2	0.2	-	0.7

表：約1 km四方の人工道路網での評価実験結果

手法	移動時間	混雑時間	総合評価
最短経路	30.9分	12.5分	156.4分
提案手法	23.5分	2.2分	45.9分

適応 Thouless-Anderson-Palmer 近似による勾配法を用いた 制限ボルツマンマシンの学習

高橋茶子 (東北大) 安田 宗樹 (山形大) 田中 和之 (東北大)

目的

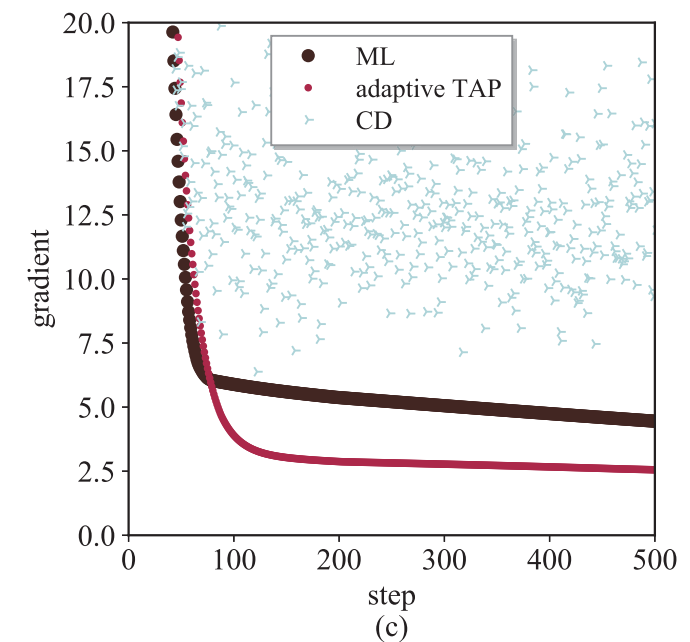
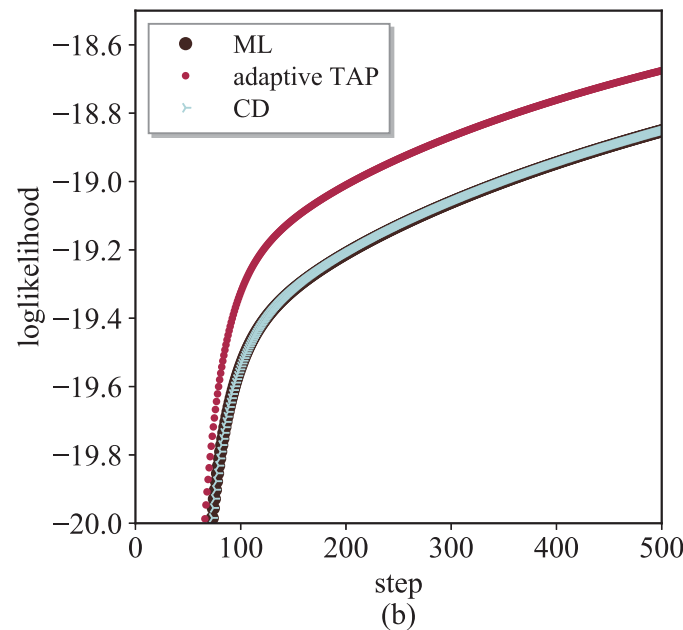
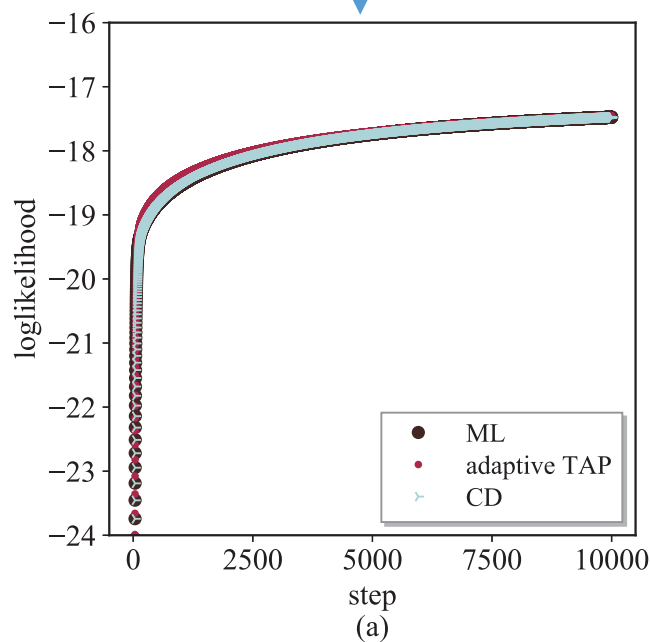
平均場近似法を用いた制限ボルツマンマシン (RBM) の学習

手法

RBM の周辺分布の期待値で勾配を書き換え → 適応 TAP 方程式で近似

結果

厳密学習・従来法と同等な性能 / 高速な尤度最大化 / 良好な勾配収束



多目的最適化と多点探索のためのベイズ最適化

和田 堯（神戸製鋼所） 日野 英逸（統計数理研）

目的 「複数」の未知関数の「効率的」な最適化 ⇒ 多目的最適化 × 多点探索

手法 多目的最適化 × 多点探索を扱うことが可能なベイズ最適化を提案

結果 テスト関数にて提案手法が効率的に最適解を探索可能なことを確認

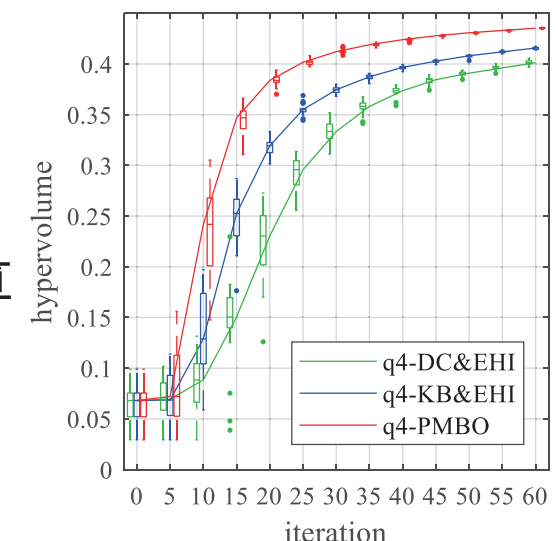
- 一般的なベイズ最適化は, 「1つの未知関数の最適化」「1点ずつ評価する最適化」
⇒ 現実には「複数未知関数の最適化ニーズ」「並列評価可能なシチュエーション多数」

- 多目的最適化 × 多点探索のベイズ最適化

Point1. 多目的最適化 × 多点探索用の新獲得関数を設計
(しかし, 単純には獲得関数の最大化困難)

Point2. 獲得関数勾配の高精度 & 低計算コストな近似計算
による効率的な獲得関数の最大化

⇒ テスト関数にてヒューリスティック探索(DC, KB&EHI)よりも
提案手法(PMBO)が効率的に探索可能なことを確認



テスト関数(DTLZ2), 1度に4点探索

批評器の複雑度はWGANの性能にどのように影響を与えるか?

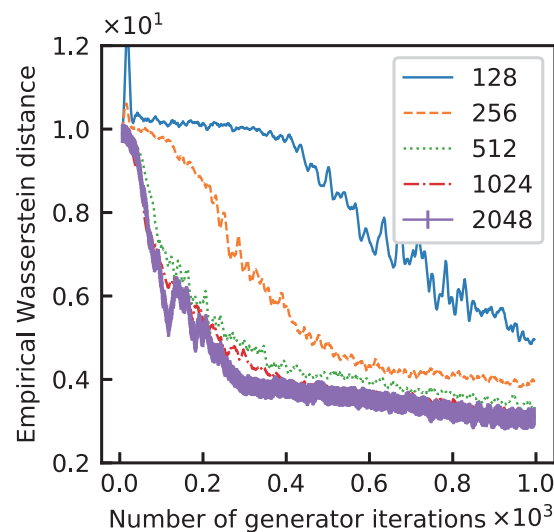
庵原明洋(京大情報学) 田中利幸(京大情報学)

目的 批評器の複雑度がWGANの性能に与える影響を分析する

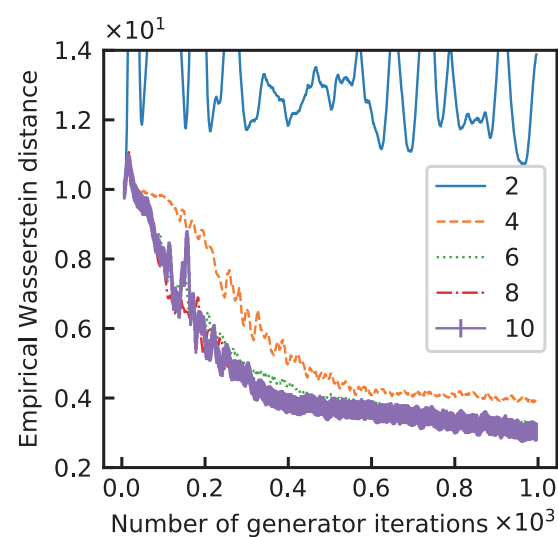
手法 批評器のみの複雑度を3つの指標で変化させてWGANの性能を評価する

結果 どの観点でも複雑度を増加させればWasserstein距離の改善に貢献した

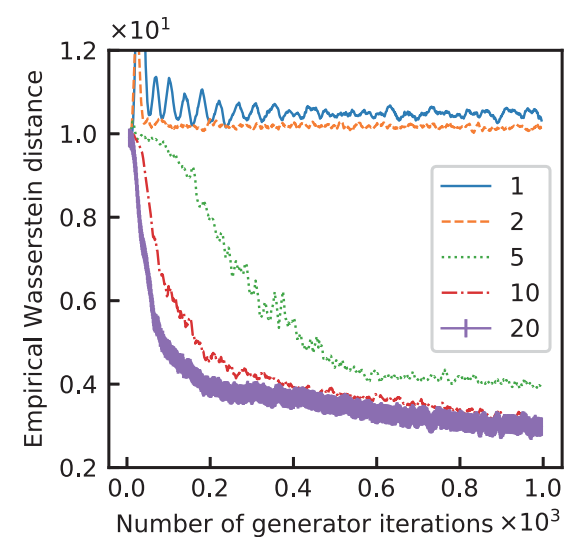
① ネットワークの幅



② ネットワークの深さ



③ 批評器の更新回数



また, インセプションモデルによる特徴空間上での経験Wasserstein-2距離が
フレッシュインセプション距離(FID)に代わる有望な評価指標であることを示唆

ミニマックス戦略による 変動ベルヌーイ過程のオンライン予測

小永吉 健太(豊橋技科大) 渡辺 一帆(豊橋技科大)

目的 変動ベルヌーイ過程の最適なオンライン予測を行う手法を開発

手法 リグレットのミニマックス問題の解析, 変分近似の適用

結果 新たな手法を開発し, 他手法よりもリグレットが少なくなることを確認

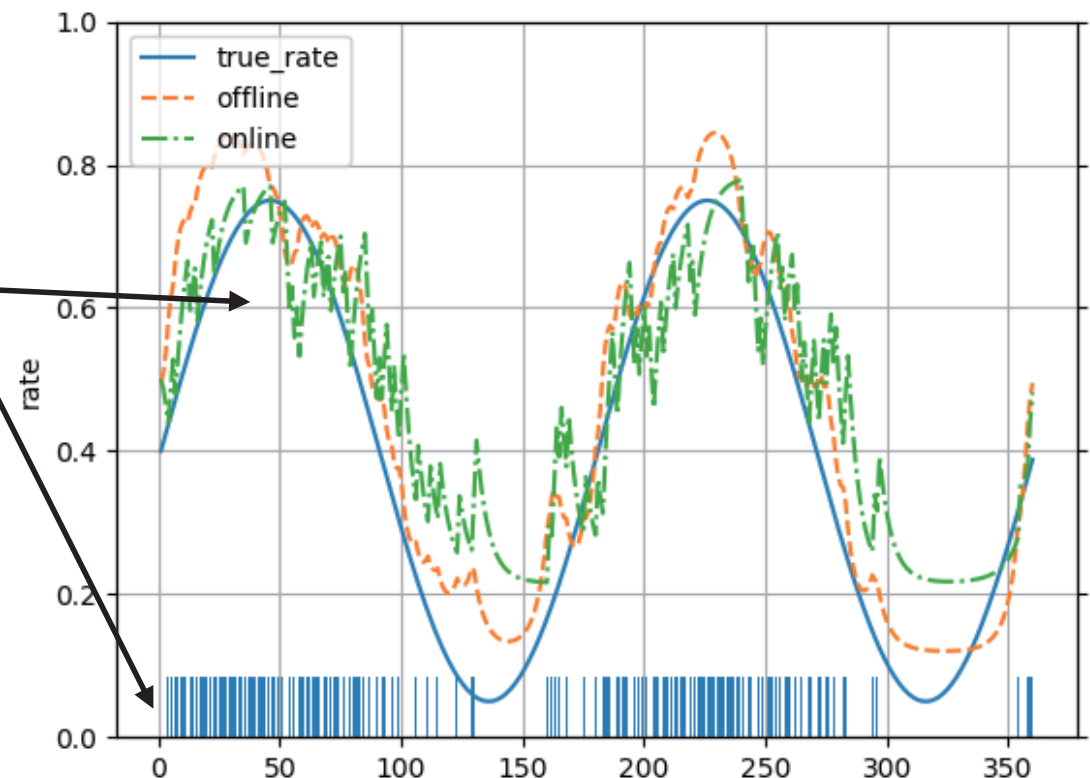
問題概要

- ① 0,1の二値系列が与えられる
- ② 逐次, 各時点で
1が出る確率を予測したい

提案手法概要

- ・ リグレットの
ミニマックス問題に着目
- ・ 損失関数に対する変分近似の適用

最適性をもったオンライン予測を実現



非定常多腕バンディットアルゴリズムを用いた ハイパーパラメータ最適化フレームワークの提案

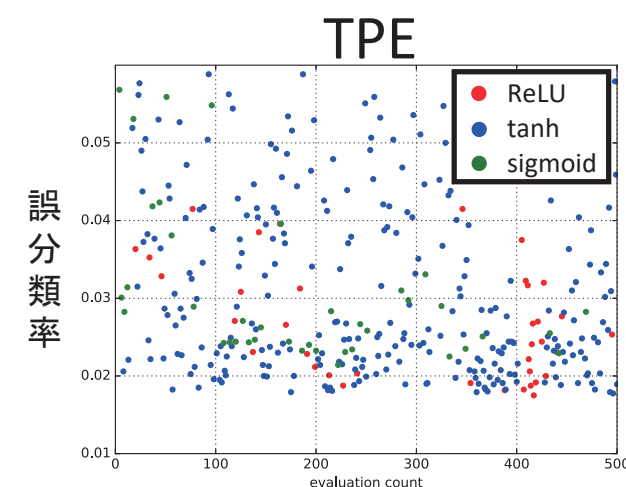
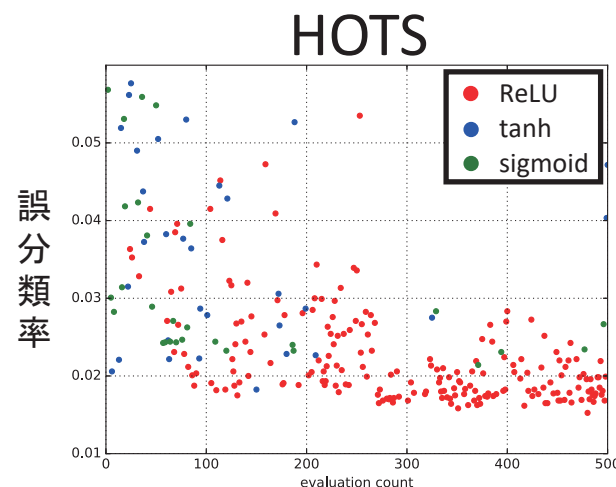
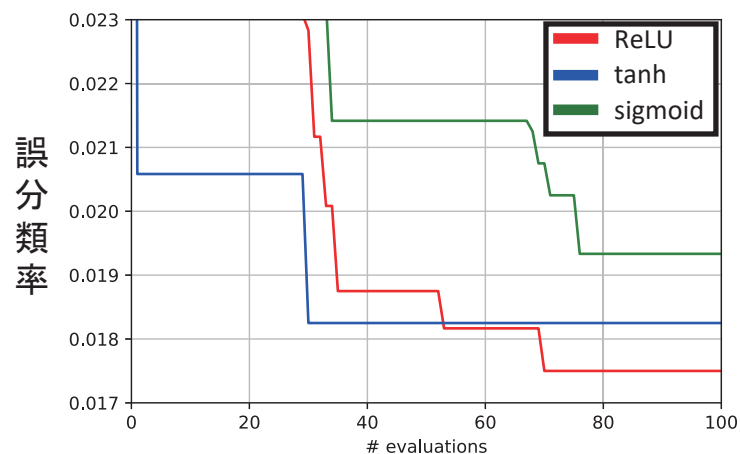
阿部 拳之 (CyberAgent, Inc.) 野村 将寛 (CyberAgent, Inc.)

目的 ハイパーパラメータ最適化において、より良い質的変数選択を実現する

手法 非定常Thompson Samplingを用いて状況に適した質的変数を選択する
ハイパーパラメータ最適化フレームワークHOTSを提案

結果 既存手法が失敗するケースにおいてもHOTSは成功することを確認

- 活性化関数としてReLUが最適となる問題
 - HOTS : ReLUを重点的に選択するため成功
 - TPE : tanhを重点的に選択するため失敗



パラメータの指数重み付けによる 深層学習モデルへの電子透かしの埋め込み

南波 涼太 (筑波大) 佐久間 淳 (筑波大・理研AIP・JST CREST)

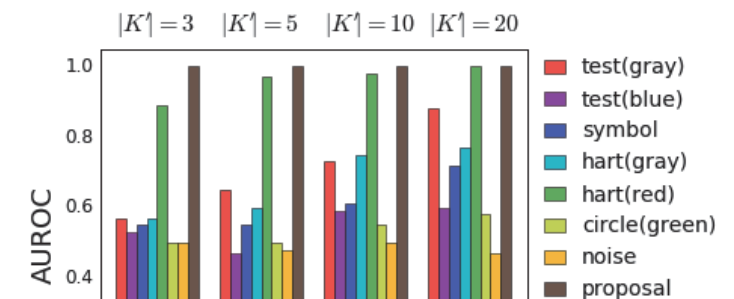
目的 電子透かしにより外部から深層学習モデルの所有権を検証する

手法 パラメータの指数重み付けを行い透かしを埋め込む

結果 検証妨害下でも既存手法に比べて高い検証性能を達成した

- 深層学習モデルの学習にはコストがかかる (e.g. データ準備や学習時間)
- モデルそのものが価値のあるもの
- サービス提供においてモデル流出の対策が必要
- 電子透かしを利用して所有権を外部から検証する
- 本研究の貢献
 - 既存手法がクエリ修正という妨害によって検証性能が大幅に低下することを実験的に示した
 - クエリ修正という妨害に耐性のある電子透かしを提案した

検証性能をAUROCで評価
各集団で最右が提案手法
で最も良い性能



部分グラフに基づくグラフ間の距離学習

吉田 知貴(名工大) 竹内 一郎(名工大) 鳥山 昌幸(名工大)

目的

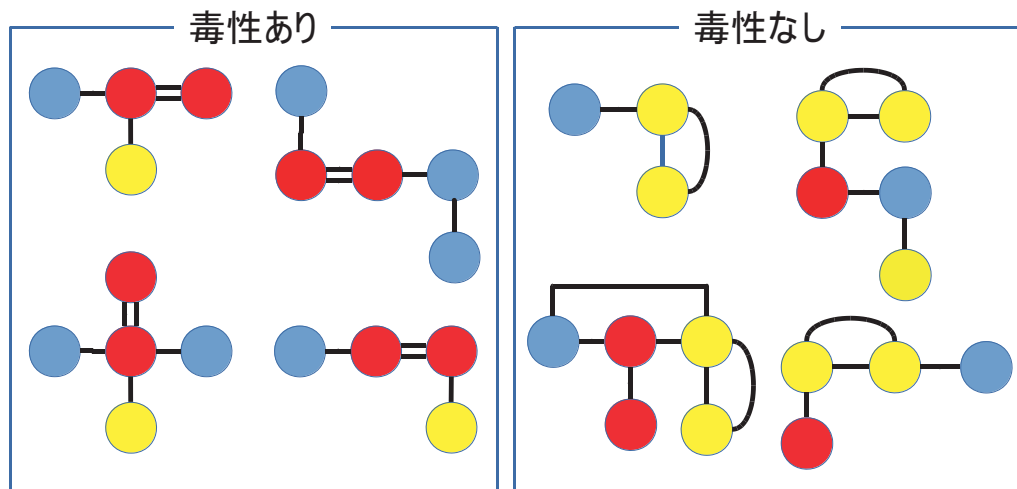
解釈性の高いグラフ間距離尺度を教師つき学習で獲得すること

手法

全ての部分グラフに対する重みの最適化を,
Safe ScreeningとWorking Set Selectionの枝刈り拡張で実現

結果

最適性を保証しつつ,
膨大な候補中の重要部分グラフのみに基づく距離尺度の獲得が可能に



既存のグラフ間の距離尺度

- Graph Kernel
 - Random Walk
 - WL Kernel
 - Deep Graph Kernel
 - ...
 - 編集距離
 - Graph embedding
 - graph2vec
- 教師なし学習
or
解釈性が低い

提案法

- 部分グラフに基づいた距離尺度を教師ありで学習
- 重要部分グラフを発見可能
- Safe Screening & Working Set Selectionと Graph Miningによる計算効率化

Discovery of Governing equations in Reproducing Kernel Hilbert space

Yosuke Otsubo (Nikon) Shinichi Nakajima (TUB · RIKEN AIP)

目的 時系列データから支配方程式を知りたい

手法 スパース推定とグラディエントマッチングを合わせた推定手法を提案

結果 シミュレーションを行い、従来法よりも高精度な推定結果を得た

PROBLEM SETTING

Time series data

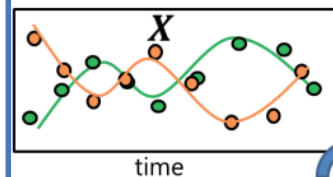
Governing Equations

$$\dot{x} = \frac{dx}{dt} = f(x; \theta)$$

$$\begin{aligned}\dot{x}_1 &= -3x_1x_2 + 0.5x_3 \\ \dot{x}_2 &= -3x_1x_2 + 2x_3 \\ \dot{x}_3 &= 3x_1x_2 - 2x_3 \\ \dot{x}_4 &= 1.5x_3\end{aligned}$$

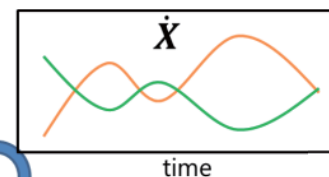
kernel regression

$$x \sim g(\alpha)$$



derivatives

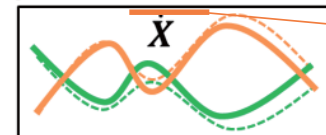
$$\dot{g}(\alpha)$$



interpolant

sparse inference

$$\dot{g}(\alpha) \sim \Theta\beta + \lambda_1 \|\beta\|_1$$



gradient matching
with feature selection

CONCEPT of METHOD

$$\begin{array}{c} \begin{array}{|c|c|c|} \hline x & y & z \\ \hline \end{array} \\ \text{time} \end{array} = \begin{array}{c} \begin{array}{|c|c|c|c|c|c|} \hline 1 & x & y & z & x^2 & xy \\ \hline \end{array} \\ \text{time} \end{array} \dots \begin{array}{c} \begin{array}{|c|c|c|} \hline \beta_x & \beta_y & \beta_z \\ \hline \end{array} \\ \text{time} \end{array}$$

Derivatives Candidate of functions Coefficients

海底水圧データを用いた津波高予測手法の比較

柏原 健之朗（東大） 五十嵐 康彦（東大/物質・材料研究機構/JST） 吉川 真史（東大）
馬場 俊孝（徳島大） 堀 高峰（海洋研究開発機構） 岡田 真人（東大/物質・材料研究機構）

目的

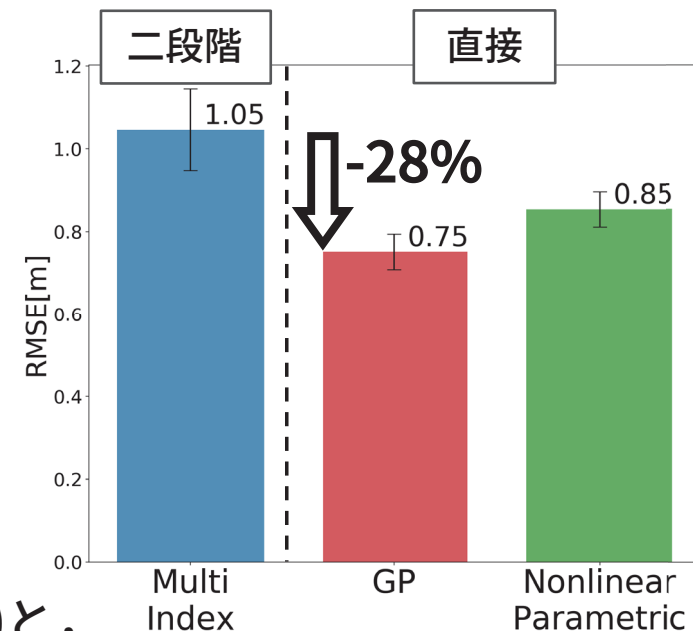
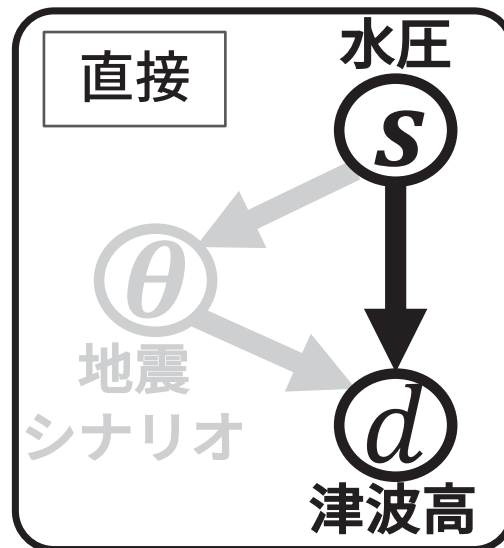
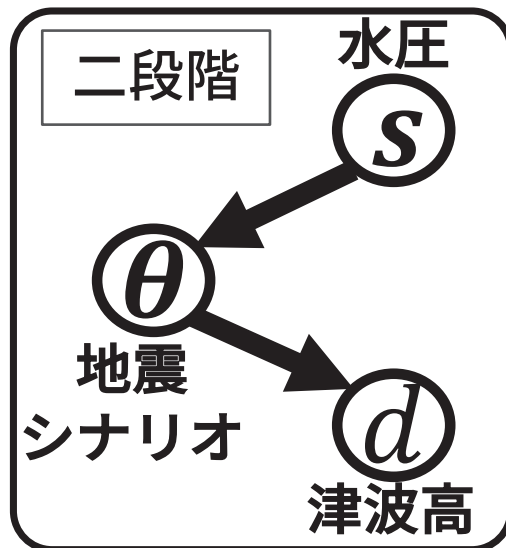
海底水圧データを用いた津波高即時予測における二段階アプローチと機械学習による直接アプローチの予測性能の比較・検討

手法

南海トラフ地震を想定した津波シミュレーションデータを用いた沿岸津波高の予測検証

結果

二段階アプローチに比べて機械学習による直接アプローチの方が予測誤差は小さく、即時性にも優れ、外挿予測誤差も小さい。



予測誤差の比較

地震シナリオ推定を経由する二段階アプローチ(左)と、水圧と津波高の相関を利用する直接アプローチ(右)

再生核ヒルベルト空間上のPerron-Frobenius作用素を用いた力学系間の比較について

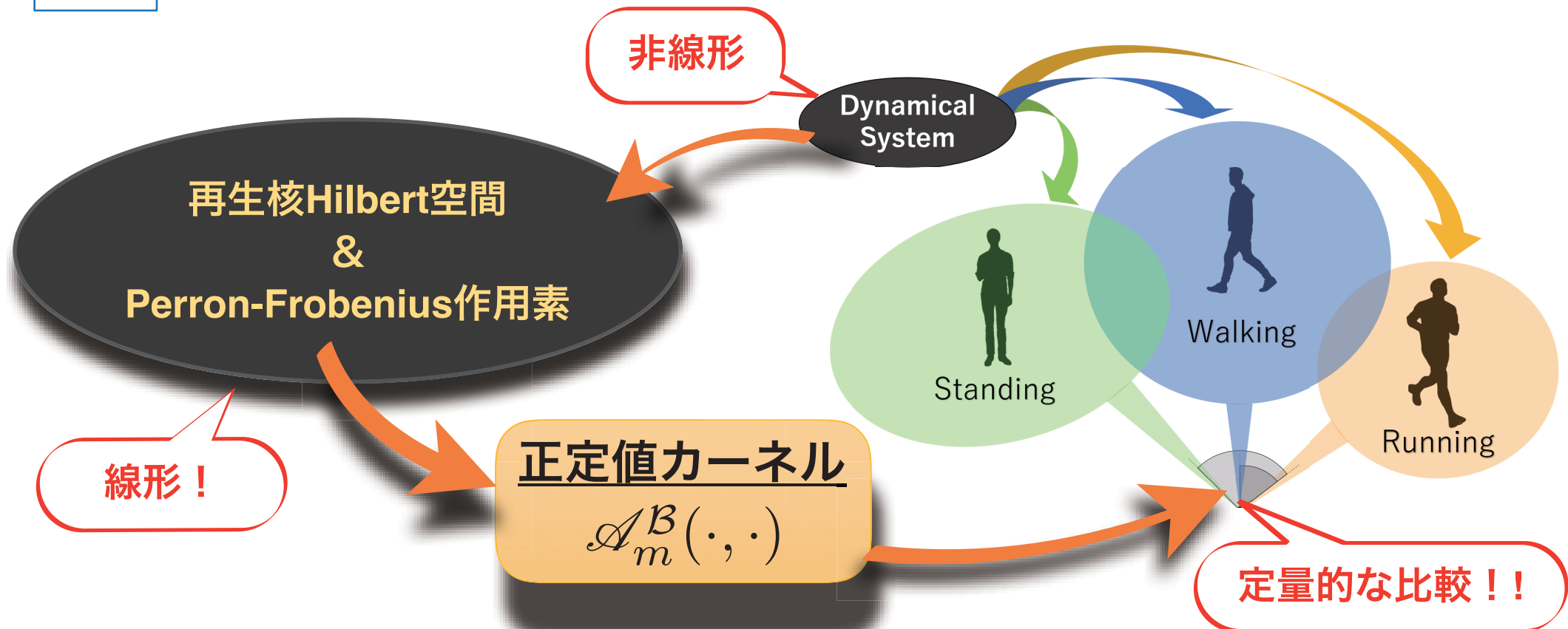
石川 勲（理研/慶大） 藤井慶輔（理研）

池田正弘（理研/慶大） 橋本悠香（慶大） 河原吉伸（阪大/理研）

目的 時系列データの分類・クラスタリング

手法 力学系モデルを仮定し，再生核Hilbert空間に定まるPerronFrobenius作用素の関数解析的考察

結果 力学系モデル間の差異を計る正定値カーネルの構成



劣モジュラ関数最大化問題に対する 効率的な分枝限定法

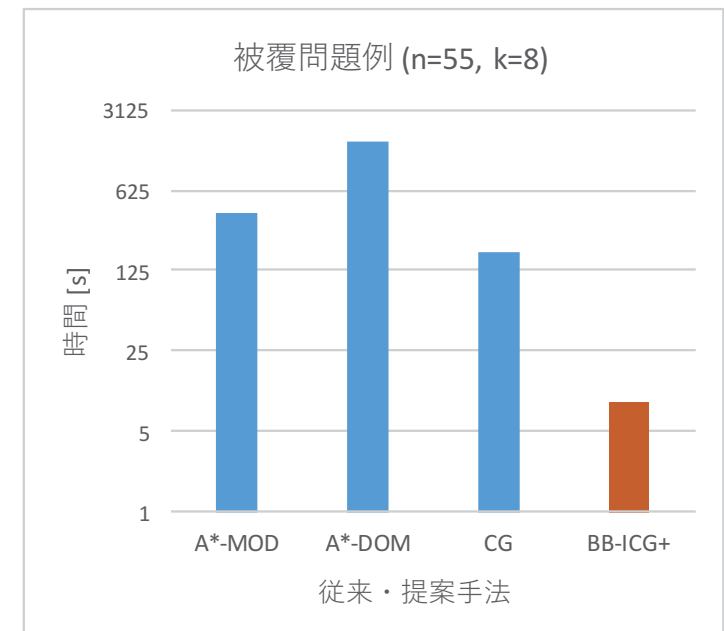
植松 直哉 梅谷 俊治 河原 吉伸

目的 劣モジュラ関数最大化問題に対して厳密解法を提案する。

手法 整数計画問題に基づく分枝限定法に提案する。

結果 ベンチマーク問題例において提案手法が従来手法より良い性能を示した。

- NemhauserとWolseyは、劣モジュラ最大化問題を膨大な制約式をもつ整数計画問題に定式化した。
- 少数の制約式をもつ整数計画問題から始めて、上界を改善する制約式の集合を逐次追加する改良制約生成法を提案した。
- 改良制約生成法を分枝限定法に組み込んだ。
- 従来手法と比べて、計算効率を改善した。



ガウスマルコフ確率場モデルを用いた格子モデル選択

伊藤 浩理 (東大) 坂本 浩隆 (東大/学振)
片上 舜 (東大/学振) 岡田 真人 (東大/NIMS)

目的

人工的に生成した 1 次元画像の真の格子モデルの推定

手法

二種類の格子モデルに従う一次元画像を作成

ガウスマルコフ確率場(GMRF)モデルによって, 画像の従う格子モデルを選択

結果

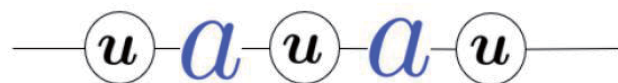
パラメータ依存性はあるが, 格子モデル選択は可能

(A)人工データ生成に用いた
2種類の格子モデル

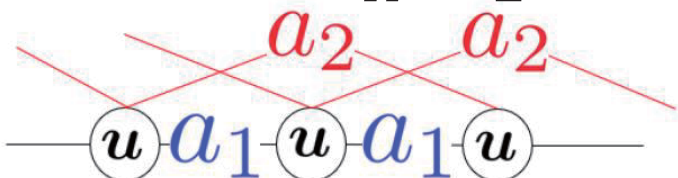
(B)GMRFモデルによる
モデルの事後確率

(C)ノイズ強度と図(A)の a_2 を
変化させた際の $p(K = 2|v)$

第1近接相互作用のみ $K = 1$



第2近接相互作用含む $K = 2$



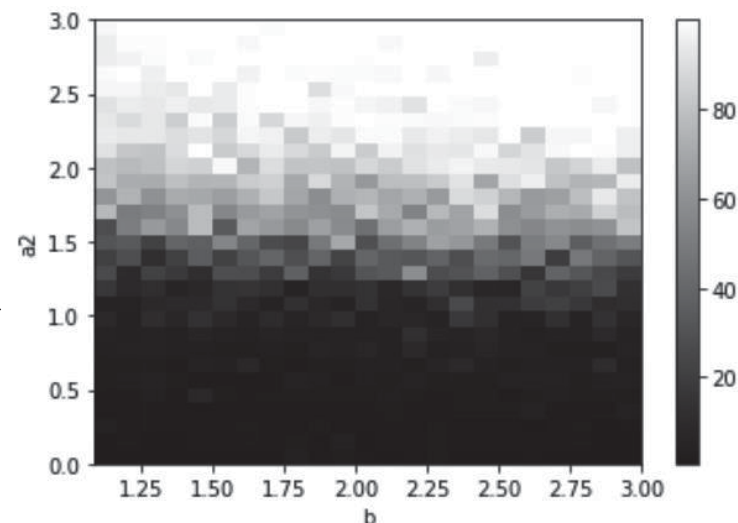
$$p(k|v) = \frac{p(v|k)p(k)}{p(v)}$$

K : モデルインデックス

v : 人工データ

a_1 : 第一近接相互作用

a_2 : 第二近接相互作用



b : ノイズ強度

階層型クラスタリングに基づく 特徴選択のためのSelective Inference

鈴木 健太 (名工大) 井上 茂乗 (名工大) 梅津 佑太 (名工大)
竹内 一郎 (名工大・理研・NIMS)

目的

階層型クラスタリングで得られる
クラスタを特徴づける変数を同定する

手法

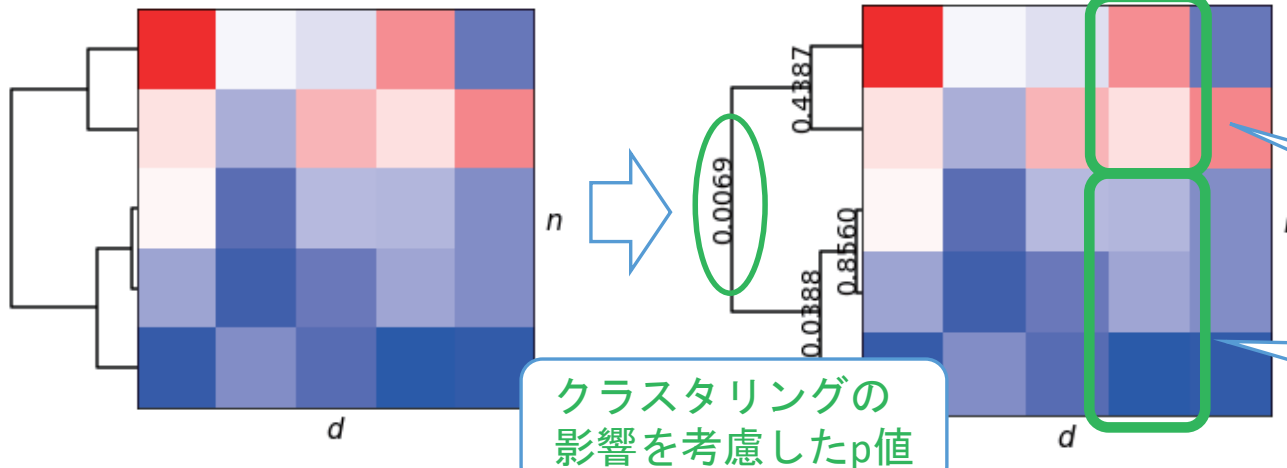
Selective Inferenceによるクラスタリング
の影響を考慮した検定を行う

結果

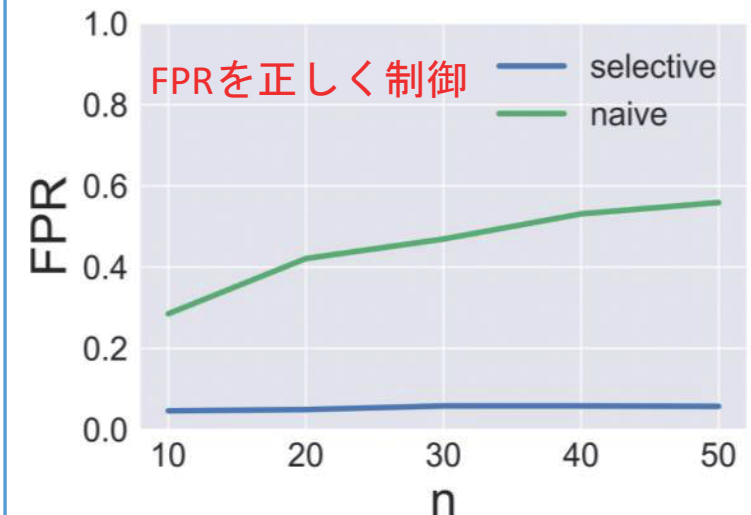
Selective Inferenceの意味でのp値が
付与された樹形図が得られる

ウォード法を適用し、
クラスタの階層構造を得る

各変数でクラスタ中心間の差
が有意かどうか検定する



実験結果



selective : 提案手法

naive : クラスタリングの影響を
考慮しない検定

非周期的データに対するスパース基底表現の推定

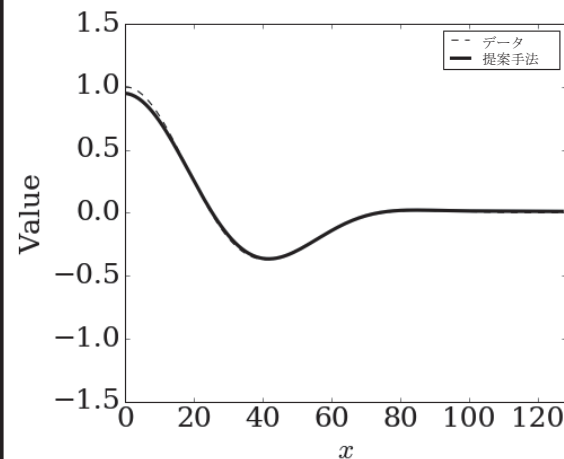
片上 舜 (東大・学振) 坂本 浩隆 (東大・学振)
五十嵐 康彦 (東大・さがけ) 岡田 真人 (東大)

目的 非周期的なデータに対するスパースな基底表現を推定する手法提案。

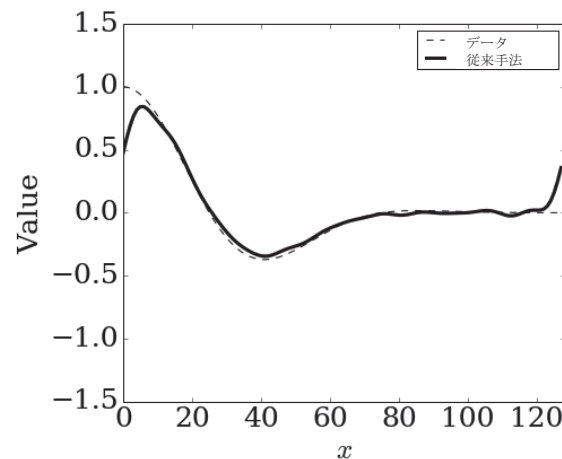
手法 データを偶関数拡張し，Fourier基底に対してLASSOにより推定する。

結果 非周期的データを表現する少数の重要な基底推定に有効性がある。

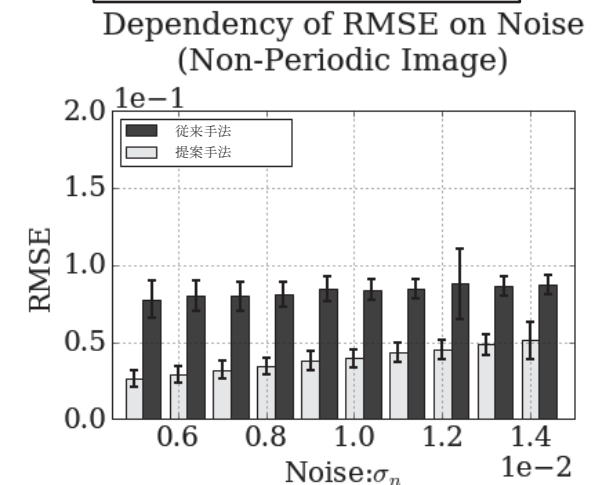
提案手法



従来手法



RMSE



移動系列データにおける特徴的なパターンを抽出するための統計的検定

佐久間 拓人(名工大) 竹内 一郎(名工大/物質・材料研究機構/理研)

目的

移動系列データからラベルに特徴的なパターンの抽出

手法

多重検定補正をした統計的検定によってラベル間の出現数の差を検定

結果

Westfall-Young法よりも高速にパターンの抽出に成功

ウミドリのおス／メス間の軌跡の違いを統計的検定によって検出



- 長大な系列であっても探索的に抽出可能
- 比較手法よりも200倍以上高速

※図中の赤い部分がそれぞれのラベルにおいて特徴的なパターン

反復識別法を用いた臨床データのミスラベル探索

藤田 雄一郎（島津製作所） 野田 陽（島津製作所） 吉村 健太郎（山梨大）
竹田 扇（山梨大） 梶原 茂樹（島津製作所）

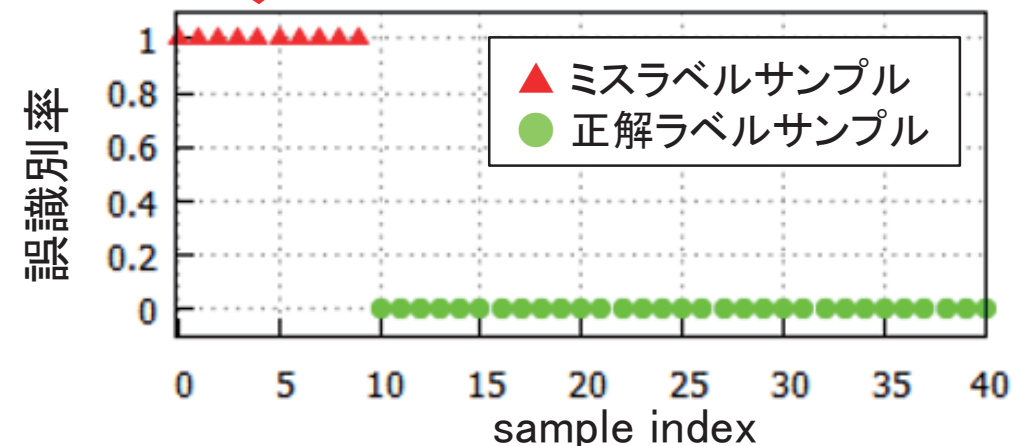
目的 学習データに含まれるミスラベルサンプルを探索

手法 複数の識別モデルを適用し、誤識別率が高ければミスラベルサンプルと判定

結果 シミュレーションデータ、ヒト肝臓がん臨床データで有効性を確認

- ID (Iterative Discrimination) 法と命名
 - SVMとRandom Forestで評価
- ⇒ Random Forestの方がミスラベルサンプルの「誤検出」が少なく、ID法に適している

ミスラベルサンプルの
誤識別率は1



ID法 (Random Forest利用) をシミュレーション
データへ適用した結果

行列分解問題の変分ベイズ解のダイナミクス解析

玉井 智貴(茨城大) 竹田 晃人(茨城大)

目的

行列分解モデルの変分ベイズ解をダイナミクス解析する

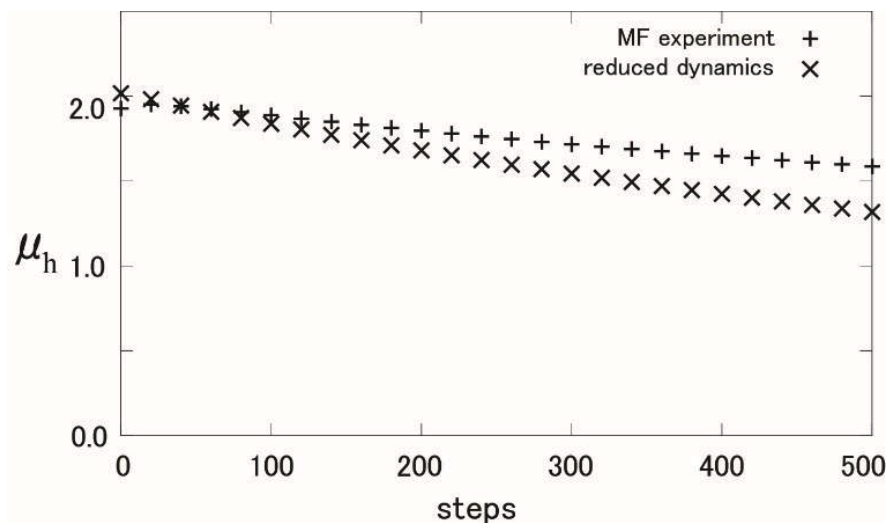
手法

ニューラルネットワークの信号雑音分離法を援用

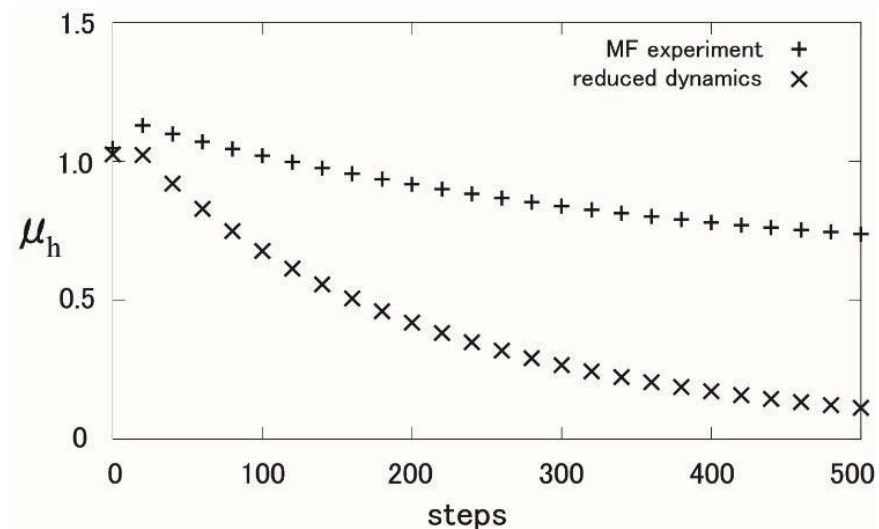
結果

ノイズ分散 σ^2 が小さいとき解析したダイナミクスは実際の挙動に近い

- 変分ベイズ解(Nakajima, Sugiyama:2011)を信号部分(μ_h, v_h)と雑音部分(干渉ノイズ)に分離



$\sigma = 0.01$ の結果



$\sigma = 0.03$ の結果

解析の詳細などはポスターへ

Cost-Sensitive MedLDAによるトラウマ患者の転帰予測

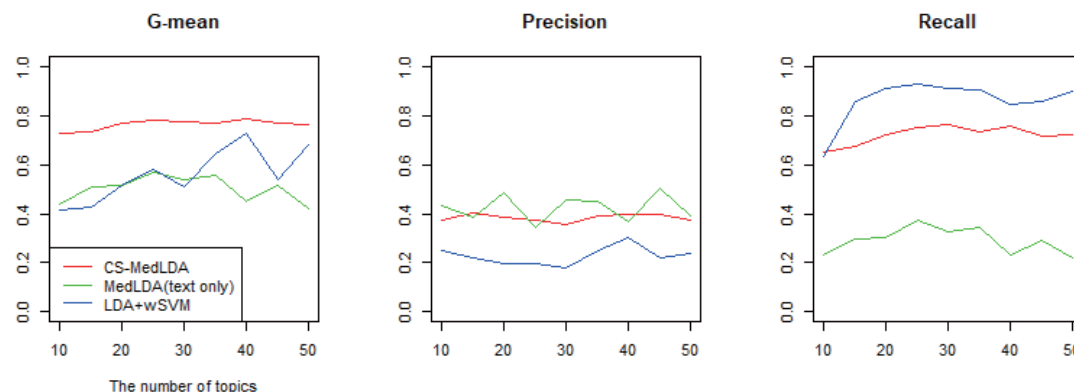
石塚 治也 石垣 司 小林 直也 工藤 大介 中川 敦寛

目的 不均衡データ判別を目的とする教師有りトピックモデルの開発

手法 コスト考慮学習を利用する教師有りトピックモデル

結果 転帰予測実験で，手法の有効性を確認

- ICUに入室した患者の院内死亡率は10%程度であり，不均衡データである．
- 不均衡データ判別を目的とした教師有りトピックモデルとして，
Cost-Sensitive MedLDAとその学習アルゴリズムを提案．
- 実証実験により，提案手法の転帰予測における有効性を示した．



実証実験結果

赤線が提案モデルによる予測の
評価結果を示す。

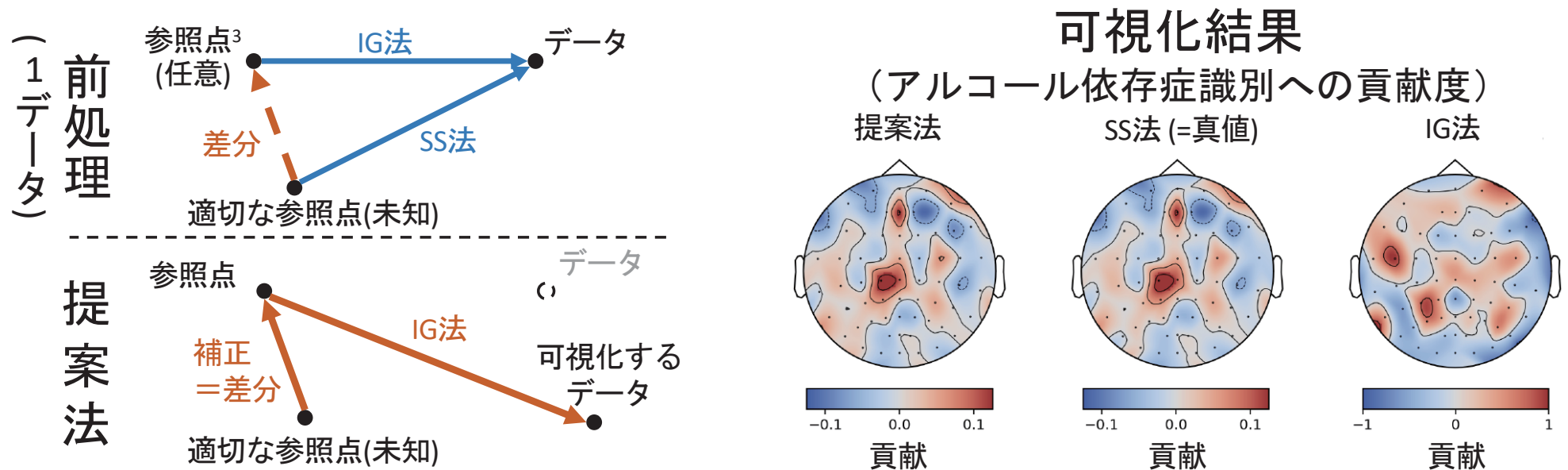
Integrated Gradientsの補正による脳波識別に貢献する特徴の可視化

立川 和樹 (阪大) 河合 祐司 (阪大) 朴 志勲 (阪大) 浅田 稔 (阪大)

目的 計算量が少なく，信頼性の高い，識別に貢献する特徴の可視化法の提案

手法 1データに対するSS法¹による貢献を用いてIG法²の貢献を補正

結果 SS法より少計算量，IG法より高信頼性であることを脳波識別で確認



1. IG法 : Integrated Gradients法 (計算量少, 不適切な参照点³の設定で, 信頼性減)
2. SS法 : Shapley Sampling法 (計算量大, 貢献の真値を計算可能 (信頼性高)) [Sundararajan et al, 2017]
3. 参照点: ハイパーパラメータ (全ての特徴を含まない入力, 適切な設定は困難) [Štrumbelj & Kononenko, 2014]

確率モデルの統合による大規模なモデルの実現

–VAE, GMM, HMM, MLDAの統合モデルの実装と評価–

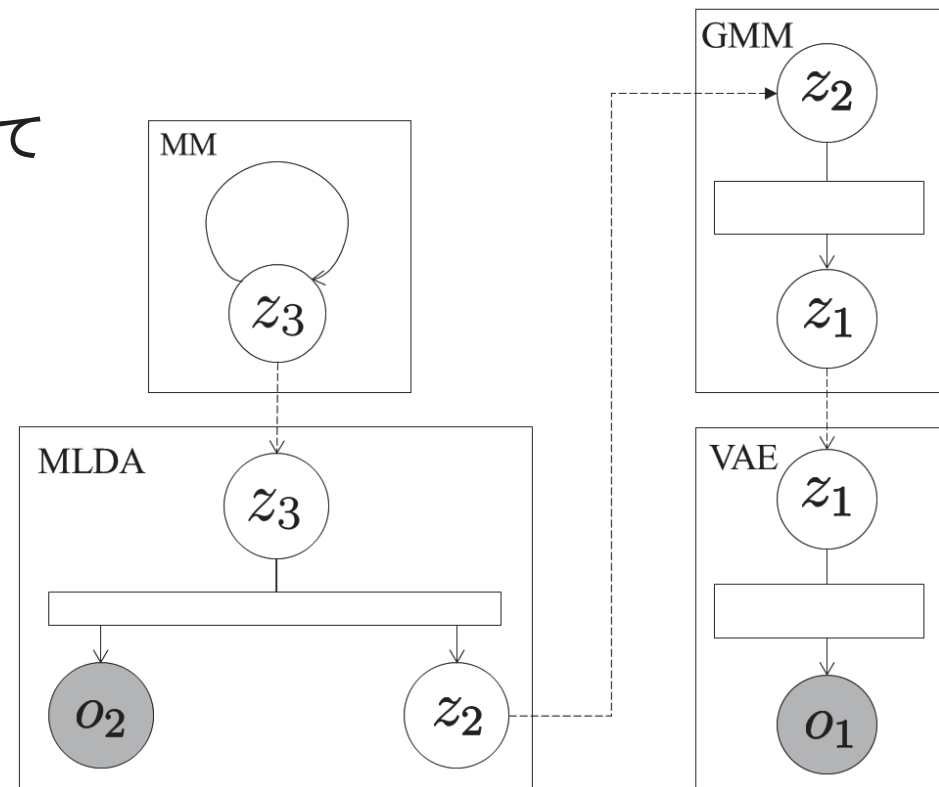
國安 瞭(電通大) 中村 友昭(電通大) 青木 達哉(電通大) 谷口 彰(立命館大) 尾崎 僚(立命館大)
伊志嶺 朝良(立命館大) 横山 裕樹(玉川大) 小椋 忠志(総研大) 長井 隆行(電通大) 谷口 忠大(立命館大)

目的 大規模な認知モデルの構築

手法 小規模なモデルを階層的に接続することによって実現

結果 Serketを用いることで容易にモデルを構築, 拡張することが可能

- Serketは小規模で基礎的なモデルであるモジュールを階層的に接続することによって大規模な認知モデルの構築を容易にするフレームワークである.
- VAE, GMM, MMモジュールを実装し, Serketを用いることで容易にそれらを統合したモデルを構築できることを示す.



Bayesian LARS-OLSによる コヒーレントフォノンの固有振動モード選択

坂田 逸志 (東大) 長野 祥大 (東大) 五十嵐 康彦 (JSTさきがけ) 村田 伸 (東大)
溝口 幸司 (阪府大) 赤井 一郎 (熊本大) 岡田 真人 (東大)

目的

格子振動の減衰過程とその他成分からなるコヒーレントフォノン信号のモード選択

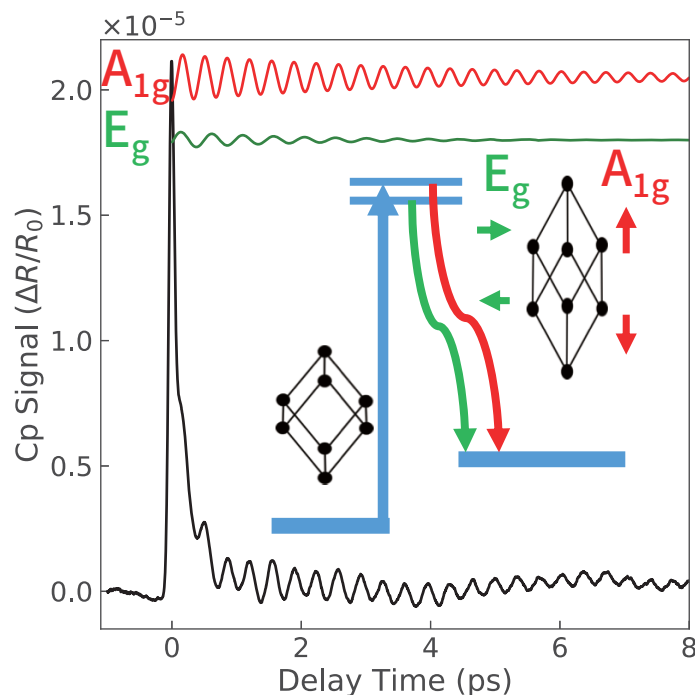
手法

動的モード分解で得られるモードをベイズ的自由エネルギーで評価

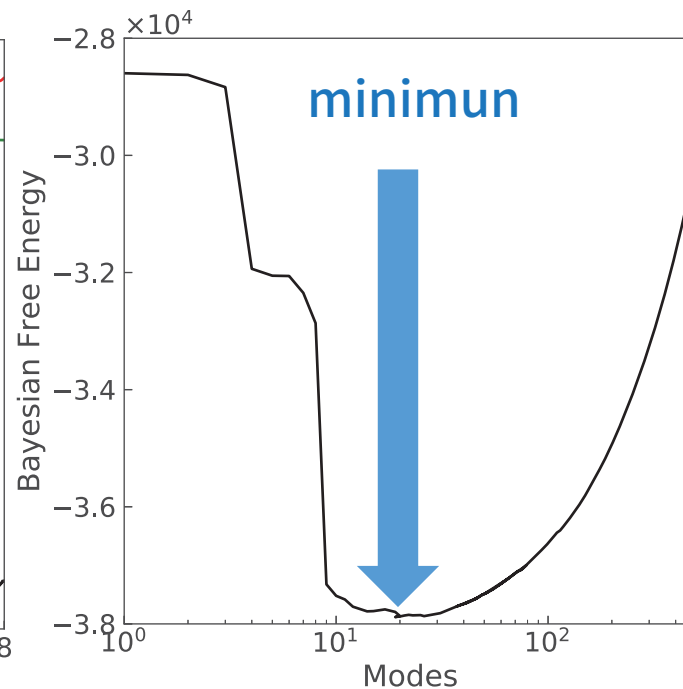
結果

推定されたモードから適切な組み合わせを選択できた

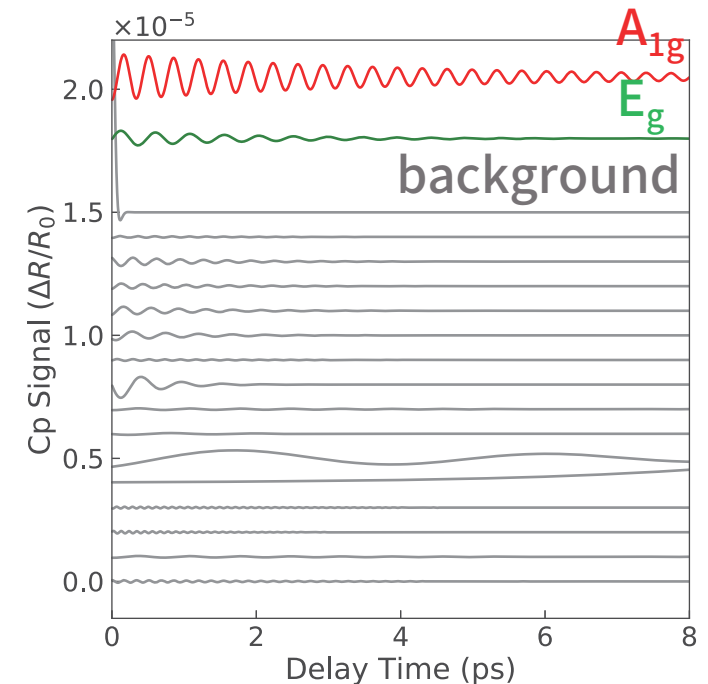
コヒーレントフォノン信号



モードの組を決定



適切なモードを選択



交換モンテカルロ法を用いた二値画像に対する ハイパーパラメータ分布推定

大日方 孝輝 (東大) 片上 舜 (東大, 学振) 楽 詠瀬 (東大) 岡田 真人 (東大)

目的

二値画像のハイパーパラメータについて信頼度を含めた評価を行う

手法

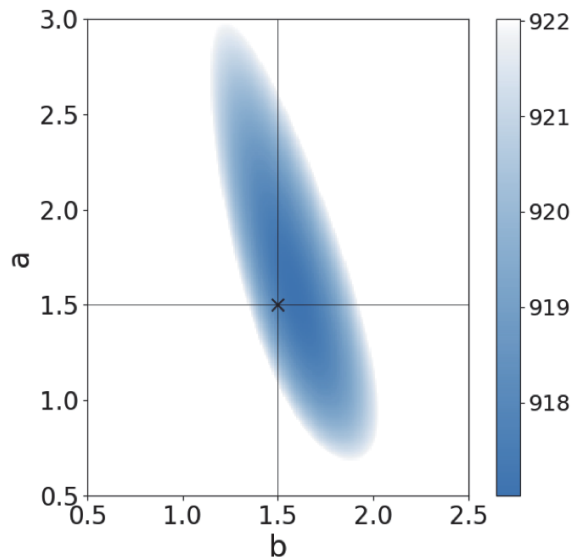
交換モンテカルロ法を用いて分配関数を計算し、事後分布を計算

結果

自由エネルギー分布から信頼度を含めてハイパーパラメータを評価可能

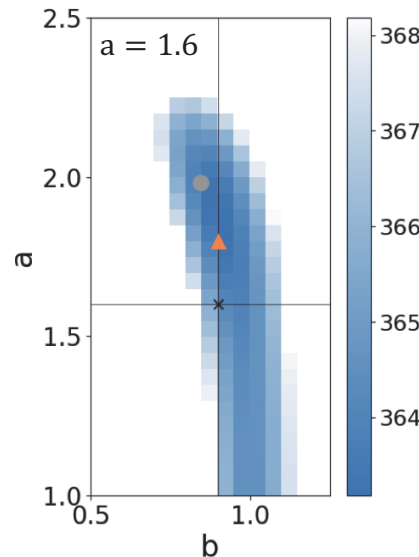
● 自由エネルギー分布図とハイパーパラメータ推定値

連続値画像

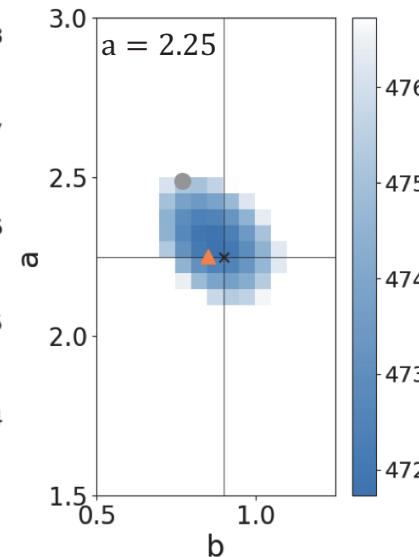


(Sakamoto et al., 2016)

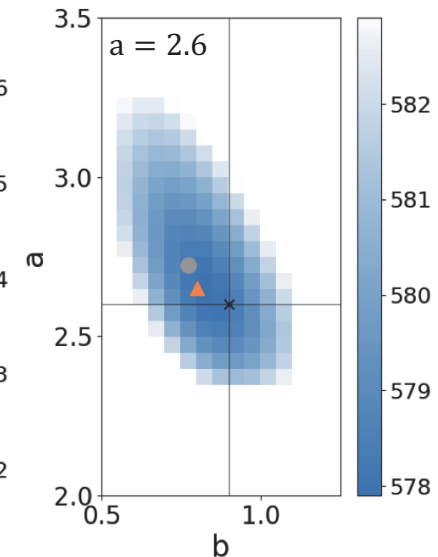
二値画像 (cf. 2次元Ising相転移点 $a_c \approx 2.27$)



推定 難



信頼度 高



信頼度 低

× : 真値

▲ : 本手法 ● : Bethe近似

ブレグマン単調作用素分解とその応用事例について

丹羽 健太(NTT MD研), Bastiaan Kleijn(Victoria University of Wellington)

目的

凸関数の和で構成されたコストに対する一般化されたソルバーの提供, 及び高速化

手法

単調作用素分解における変数空間をブレグマンダイバージェンスを用いて一般化
ニュートン法, 加速勾配法等に相当する高速収束を実現

結果

事例(全変動ノルム型ノイズ除去)に適用し, 高速/安定に収束することを確認

$$\min_{\mathbf{w}} G_1(\mathbf{w}) + G_2(\mathbf{w})$$

単調作用素分解の問題設定:
各凸コストに対して並列に最適化

(事例1) 拘束付最小化問題

$$\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{s} - \mathbf{w}\|_2^2 + \|\mathbf{D}\mathbf{w}\|_1$$

$$\min_{\mathbf{w}, \mathbf{z}} G_1(\mathbf{w}) + G_2(\mathbf{z}) \quad \text{s.t. } \mathbf{A}\mathbf{w} + \mathbf{B}\mathbf{z} = \mathbf{c}$$

双対問題に対してDouglas-Rachford型
単調作用素分解を適用 = ADMM

(事例2) 非負行列分解

$$\min_{\mathbf{W}, \mathbf{A}, \mathbf{B}} \frac{1}{2} \|\mathbf{S} - \mathbf{W}\|_2^2 + H_A(\mathbf{A}) + H_B(\mathbf{B})$$

$$\text{s.t. } \mathbf{W} = \mathbf{A}\mathbf{B}^T, \mathbf{A} \geq 0, \mathbf{B} \geq 0$$

$B(\mathbf{w} \parallel \mathbf{w}')$	$\frac{1}{2\alpha} \ \mathbf{w} - \mathbf{w}'\ _2^2$	$\frac{1}{2} \langle \mathbf{H}(\mathbf{w} - \mathbf{w}'), (\mathbf{w} - \mathbf{w}') \rangle$	
代表的な単調 作用素分解	1次勾配法 (GD)	加速勾配法 (AGD)	2次勾配法 (Newton)
Peaceman- Rachford型	✓		
Douglas- Rachford型	✓		
Forward- Backward型	✓		

動的計画法を用いた系列セグメンテーションにおける Selective Inference

戸田 博己 (名工大) 梅津 佑太 (名工大) 佐久間 拓人 (名工大)
竹内 一郎 (名工大/物質・材料研究機構/理研)

目的 選択バイアスを考慮した上で変化点の統計的有意性を適切に評価

手法 セグメント平均の差の検定をSelective Inferenceの枠組で定式化

結果 FPRを有意水準に制御し, 変化点の定量的な評価が可能に！

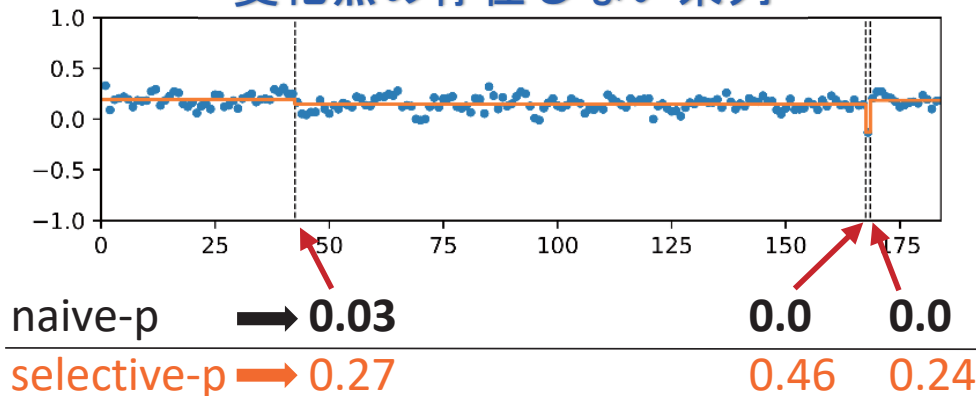
通常の検定手法 (naive)

セグメンテーション後の推論において, 検定統計量の分布に誤った分布を想定することでFPRが非常に高くなる (選択バイアス)

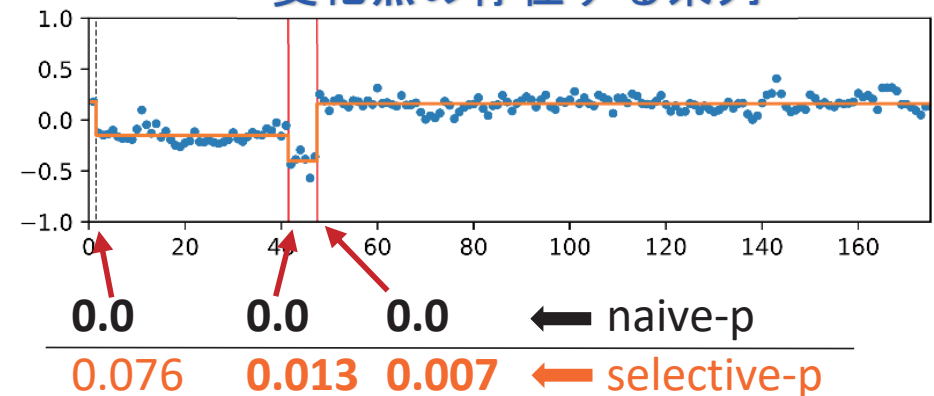
提案手法 (selective)

Selective Inferenceの枠組を用いて, 変化点が選択されたという事象を条件付けたもとで検定統計量の正しい分布を導出

変化点の存在しない系列



変化点の存在する系列



混合正規分布モデルにおけるベイズ推定と変分ベイズ推定の比較

中山 智文 藤井 直樹 (東大)

永田 賢二 (産総研/JSTさきがけ) 岡田 真人 (東大)

目的

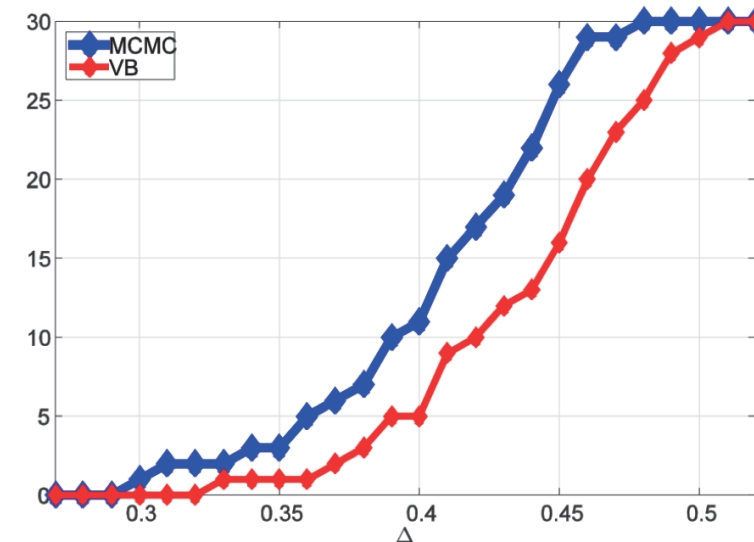
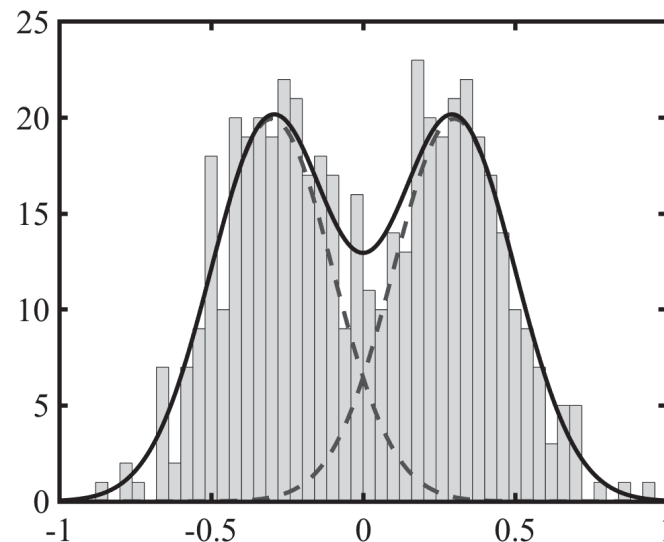
混合正規分布モデルにおけるベイズ推定と変分ベイズ推定の比較

手法

事後分布と自由エネルギーをそれぞれの手法で比較をする

結果

変分ベイズ推定はクラスタ数を少なく推定する状況があることがわかった



- 各正規分布の中心間の距離 Δ を変えながら、各 Δ に対して、1次元の混合正規分布からサンプリングした500点のデータを30セット作成
- それぞれの手法を用いてモデル選択を行い、正解した回数を調べた
- 変分ベイズ推定が、ベイズ推定よりもクラスタ数を少なく推定する領域があることがわかった

多段階学習Echo state network による 非線形時系列予測

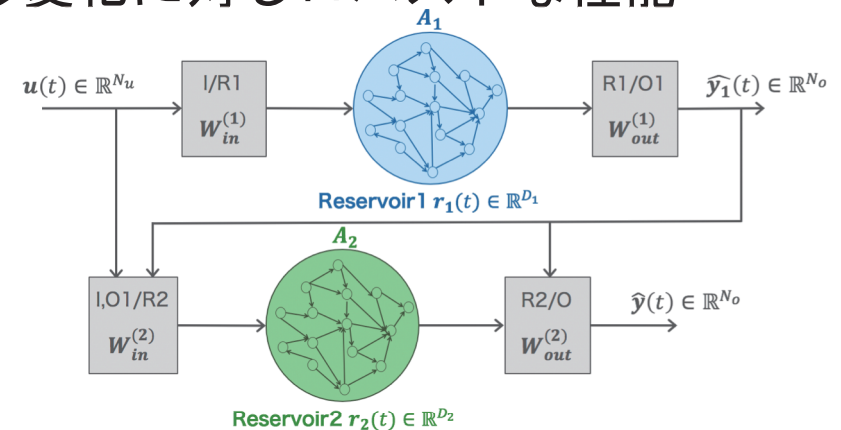
秋山 貴則 田中 剛平 (東大)

目的 リザーバーコンピューティングによる非線形時系列予測タスクの精度向上

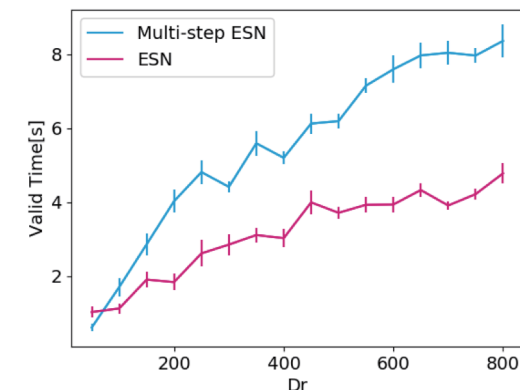
手法 Echo state network の多段階学習化による予測誤差補正

結果 予測精度改善 & ハイパーパラメータの変化に対しロバストな性能

- 従来法 [Jaeger et al.,2004] の問題点
 - リザーバーサイズに対し予測精度が頭打ち
 - ハイパーパラメータへの強い精度依存性
- 多段階学習 Echo state network



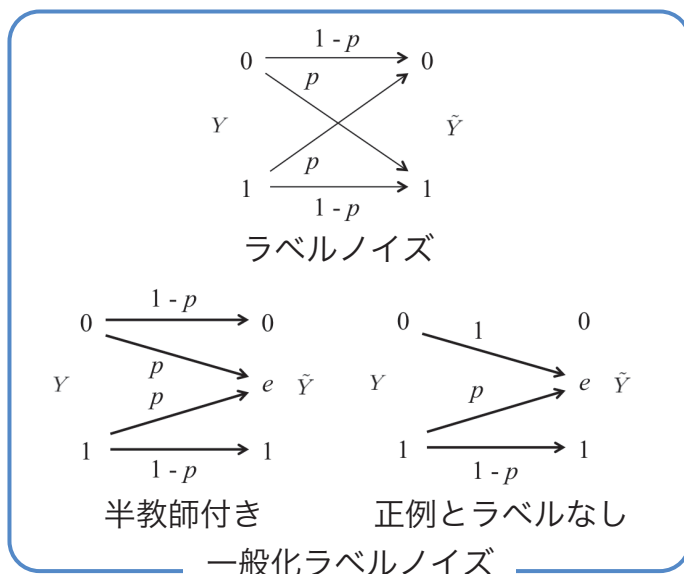
- ローレンツ方程式 & NARMAの予測精度向上
 - リアプノフ指数による評価
 - ハイパーパラメータ
 - ランダムな重み
- } に対するロバスト性



一般化ラベルノイズの下での分類に関する漸近評価

安田 豪毅 (早大) 須子 統太 (早大) 小林 学 (早大) 松嶋 敏泰 (早大)

- 目的** 一般化ラベルノイズの下での分類について理論的に性能評価
- 手法** パラメータに関する一致性を満たす尤度方程式の解に基づく分類方法
- 結果** EMアルゴリズム等の最尤法で理想的な解が得られた場合の性能を導出



$$E_{\mathbf{X}YD_n, \theta_0} \left[\log \frac{p(Y|\mathbf{X}, \theta_0)}{p(Y|\mathbf{X}, \hat{\theta}_{D_n})} \right] = \frac{1}{2n} \text{tr} \left[I_{Y|\mathbf{X}, \theta_0} I_{\mathbf{X}\tilde{Y}, \theta_0}^{-1} \right] + o\left(\frac{1}{n}\right).$$

$\hat{\theta}_{D_n}$: 一致性を満たす尤度方程式の解

$I_{Y|\mathbf{X}, \theta_0}$: $\mathbf{X}$ の下での Y の条件付きフィッシャー情報行列

$I_{\mathbf{X}\tilde{Y}, \theta_0}$: $\mathbf{X}\tilde{Y}$ のフィッシャー情報行列

Disentangledな特徴表現を用いた 意図的クラスタリングによる未知ラベルの検出

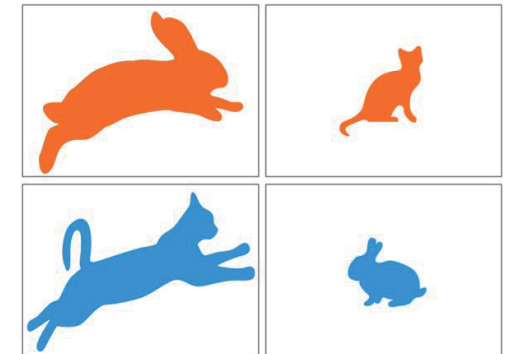
福馬 智生 (東大) 鳥海不二夫(東大)

目的 少数のラベルにより意図的にクラスタリングを促す

手法 β -VAEを用いた特徴抽出+Siamese Network

結果 意図に基づいたクラスタリングにより未知クラスの発見を行えた

- ・何に着目してクラスタリングを行うかは
何に着目するかによって異なる(色・形・大きさ・種類)
- ・少数のラベルにより、意図的にクラスタリングを促し、
その規則に基づいてもともと付与したラベルのどれにも
属さないデータの抽出を行う
 - ・ β -VAEによるDisentangledな特徴量とSiamese Networkの組み合わせにより、
上記問題が解決できることが分かった。



尤度比検定における差分プライバシー

薮島 嵐(筑波大) 佐久間 淳(筑波大・理研AIP)

目的

ロジスティック回帰における尤度比検定のプライバシー保護

手法

検定統計量に加法的なノイズを加えることによる差分プライバシーの保証

結果

検定統計量の ℓ_1 -sensitivityの導出・提案する検定の有意水準の適切な制御

- データの属性間の影響を検定したい→ロジスティック回帰における**尤度比検定**
- 検定統計量についてプライバシー保護をしたい→**差分プライバシー**
- 統計量に加法的にラプラスノイズを加えるラプラスメカニズム^[1]を利用
- 検定統計量の ℓ_1 -sensitivityを導出することで**差分プライバシー**を保証・証明
- モンテカルロ法により検定の**有意水準**を適切に制御
- **差分プライバシー**を保証した**尤度比検定**のアルゴリズムを導出
- 導出したアルゴリズムの性能評価→検定の検出力を評価

[1] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. *In Proceedings of the Third Conference on Theory of Cryptography*, TCC' 06, pp. 265–284, Berlin, Heidelberg, 2006. Springer-Verlag.

最大次数が未知の多項式回帰モデルに対する スパース推定に関する一考察

井上 一磨（早大） 清水 良太郎（早大） 須子 統太（早大） 後藤 正幸（早大）

目的

最大次数が未知の多項式回帰モデルに対して、正しくパラメータをスパース推定する

手法

モデルに有効な変数と他の変数を繰り返し乗じながら説明変数集合を増加させ、探索的に良いモデルを選択する

結果

数値実験により、提案アルゴリズムは従来手法と比べ、最大次数が未知の場合においてパラメータの推定精度が高いことを確認した

従来手法(SPORE-LASSO)

Step1:

モデルに有効な変数を選択

Step2:

Step1で得た変数を任意の d 次まで多項式展開する

Step3:

Step2で得た変数に対してパラメータを推定

提案アルゴリズム

Step1: モデルに有効な変数を選択

Step2: Step1で得た変数と元の変数集合の各要素を乗じ、新たな変数集合を取得

Step3: Step2で得た変数集合からモデルに有効な変数を選択

Step4: step3で得た変数と元の変数集合の各要素を乗じ、新たな変数集合を取得

Step5: Step2とStep4で得た変数を用いて、それぞれパラメータを推定し、平均二乗誤差(MSE)を算出



MSEの変化量により、Step3から収束まで繰り返す

実験概要: N 個の学習データを用いて、予めパラメータを設定した多項式回帰モデルのパラメータを推定する。その後、1000件のテストデータに対する平均二乗誤差を算出する。

評価指標: 平均二乗誤差

表1. 実験結果

	$N = 30$	$N = 100$
提案アルゴリズム	20.20	7.16
SPORE-LASSO($d = 2$)	30.90	26.43
SPORE-LASSO($d = 3$)	22.68	16.59

反復再重み付け最小二乗法による正則化モード回帰

山崎 遼也 (京大) 田中 利幸 (京大)

目的

正則化モード回帰のパラメータ推定に対する反復再重み付け最小二乗法 (IRLS) アルゴリズムの収束性を解析すること。

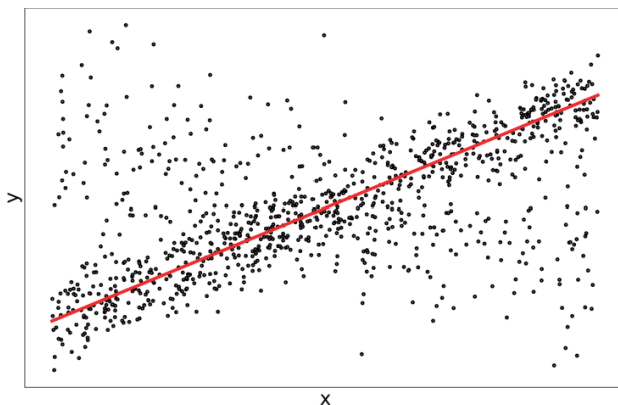
手法

下界最大化 (MM) アルゴリズムの理論, 平均値シフト (MS) アルゴリズムの解析

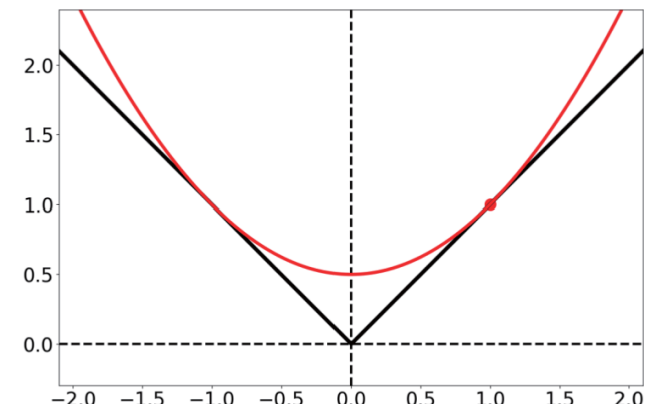
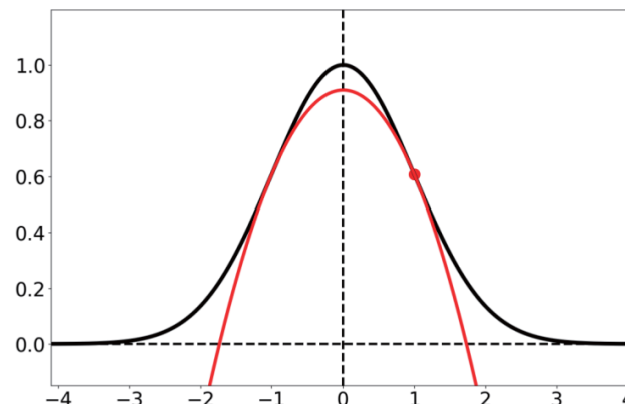
結果

- 目的関数値系列が収束するためのカーネル関数 K_h と正則化関数 P_Λ に関する十分条件を与えた。
- Epanechnikov カーネル, ℓ_2 正則化の下で目的関数値/パラメータ推定値系列が有限回の反復で収束することを証明。

正則化モード回帰の目的関数
$$O_R(\theta) = \frac{1}{n} \sum_{i=1}^n K_h(y_i - \theta^\top \phi_i) - P_\Lambda(\theta)$$



モード回帰



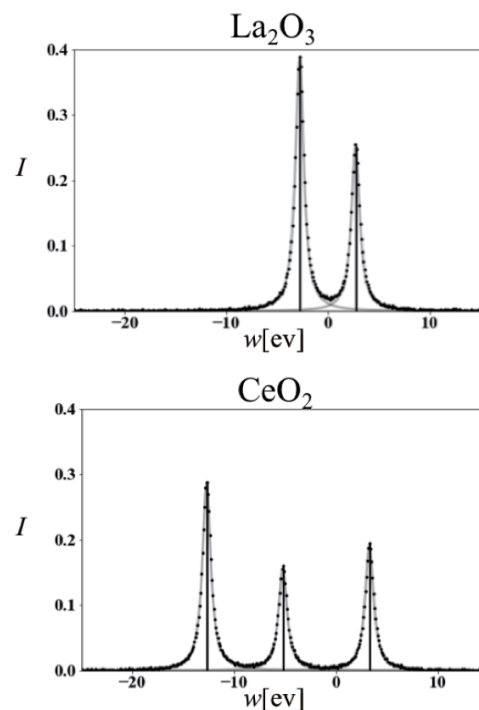
カーネル/正則化関数の最良二次下界/上界関数

X線光電子分光におけるベイズ推論によるハミルトニアン選択

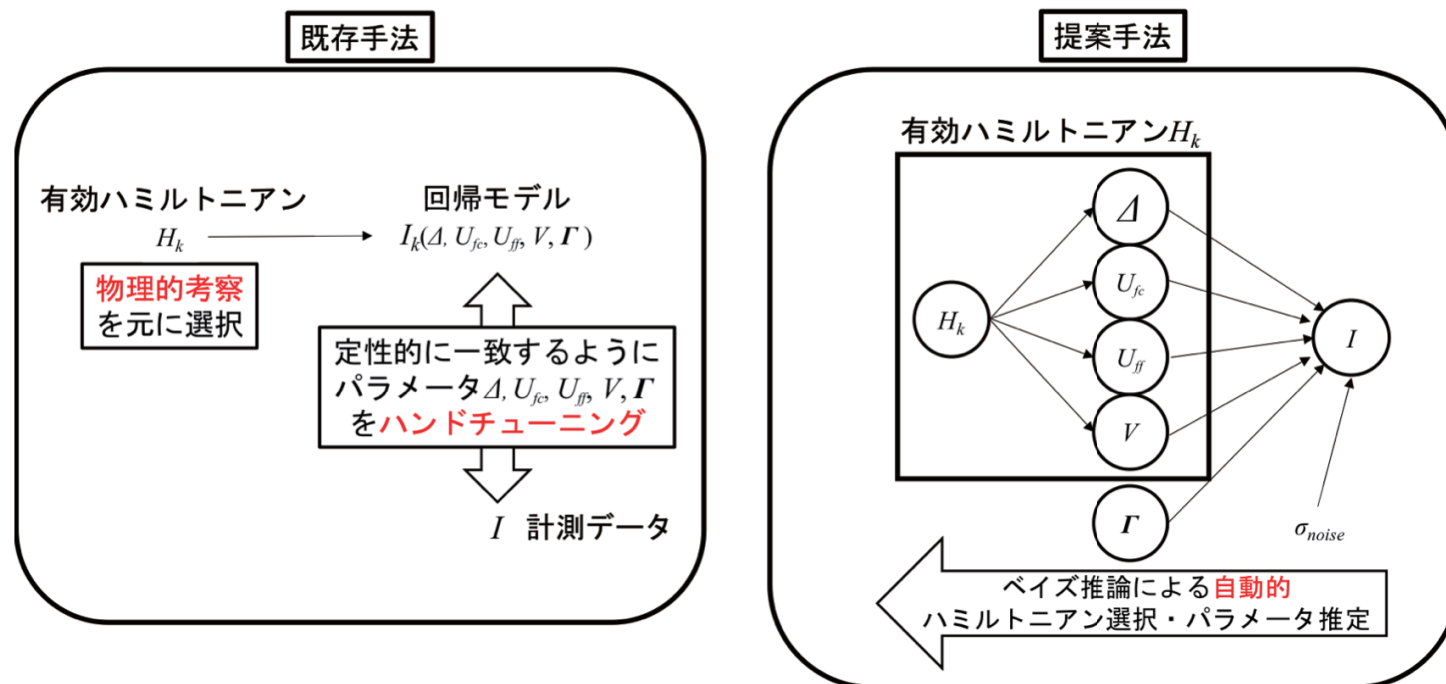
本武陽一（東大） 水牧仁一郎（高輝度光科学研究センター）
赤井一郎（熊本大） 岡田真人（東大）

- 目的** XPSスペクトルデータからの有効ハミルトニアン選択・パラメータ推定
- 手法** 交換モンテカルロ法を用いたベイズ的スペクトル分解によって実現
- 結果** 物理的知見と一致する有効ハミルトニアン選択・パラメータ推定が実現

● 分析対象のXPSスペクトル



● 既存手法と提案手法の比較



テンソル分解を用いた教師無し学習による変数選択の バイオインフォマティクスへの応用

田口善弘(中大)

目的

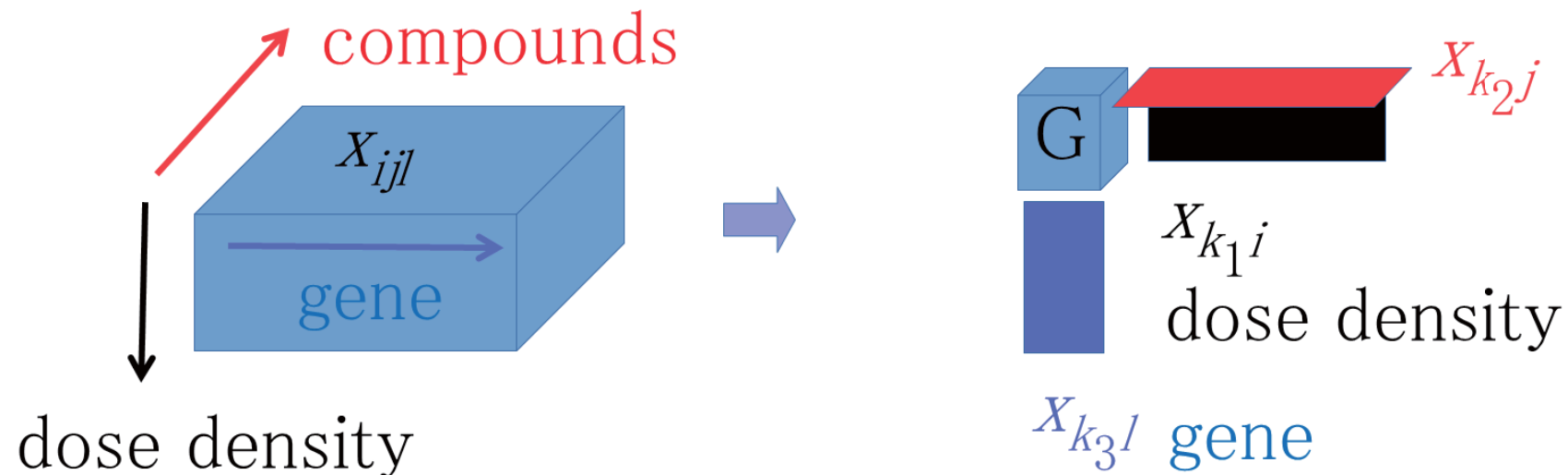
バイオインフォマティクスのlarge p small n問題における変数選択問題の解決

手法

テンソル分解を用いた教師無し学習による変数選択法

結果

数万の変数に対して数個のサンプルしかない場合でも変数選択が可能である。



$$X_{ijl} \doteq \sum_{k_1, k_2, k_3} G_{k_1, k_2, k_3} X_{k_1 i} X_{k_2 j} X_{k_3 l}$$

非対称性を利用した悪腕存在チェックアルゴリズム

田畑 公次 (北大) 中村 篤祥 (北大) 小松崎 民樹 (北大)

目的

多腕バンディット問題設定で悪腕が存在するかどうかを判定

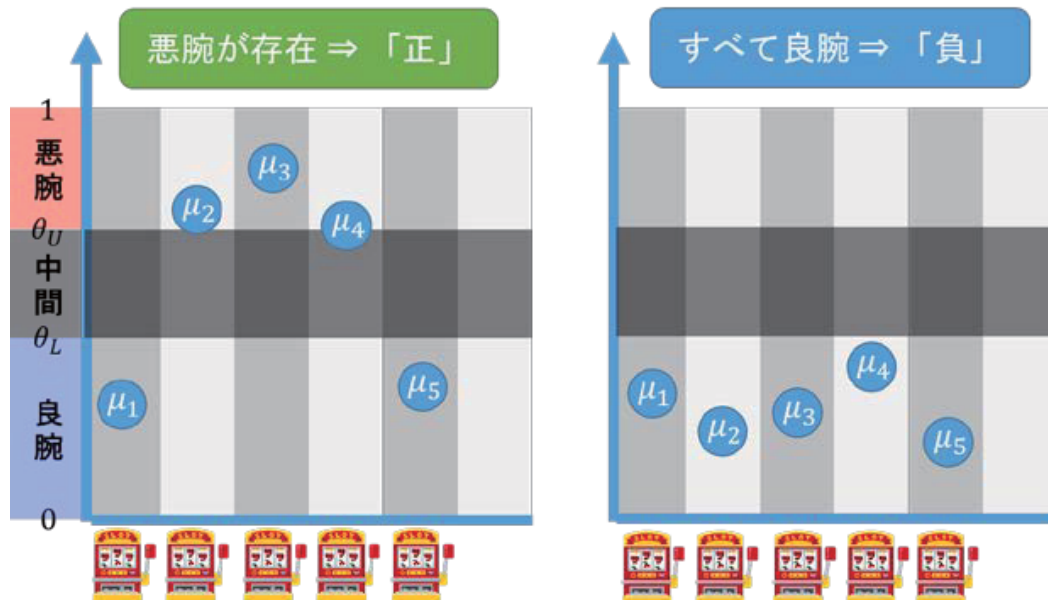
手法

非対称的な停止条件と腕選択方策の組み合わせによるアルゴリズム

結果

最悪停止時刻の理論解析とシミュレーション実験による比較

悪腕存在チェック問題



δ : 許容誤識別率

非対称性を利用した停止条件

$$\hat{\mu}_i(n) + \sqrt{\frac{1}{2n} \ln \frac{2n^2}{\delta}} \leq \theta_U \Rightarrow \text{「負」}$$

$$\hat{\mu}_i(n) - \sqrt{\frac{1}{2n} \ln \frac{2Kn^2}{\delta}} \geq \theta_L \Rightarrow \text{「正」}$$

$$n_i \hat{\mu}_i(n) - \sqrt{\frac{t}{2} \ln \frac{2t^2}{\delta}} \geq t\theta_L \Rightarrow \text{「正」}$$



腕選択方策

UCB, LUCB, modified-APT

パラメータ未知の 一般化ラベルノイズモデルにおける分類法について

須子 統太（早大） 安田 豪毅（早大） 堀井 俊佑（早大） 小林 学（早大）

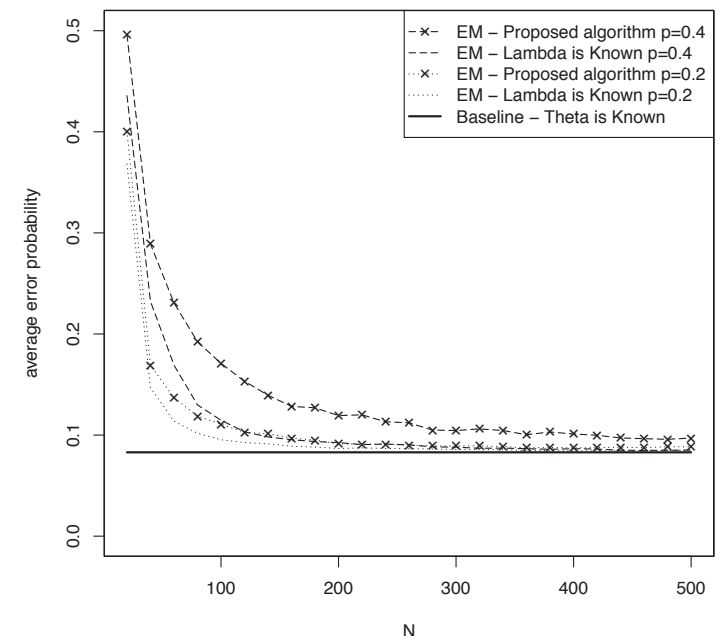
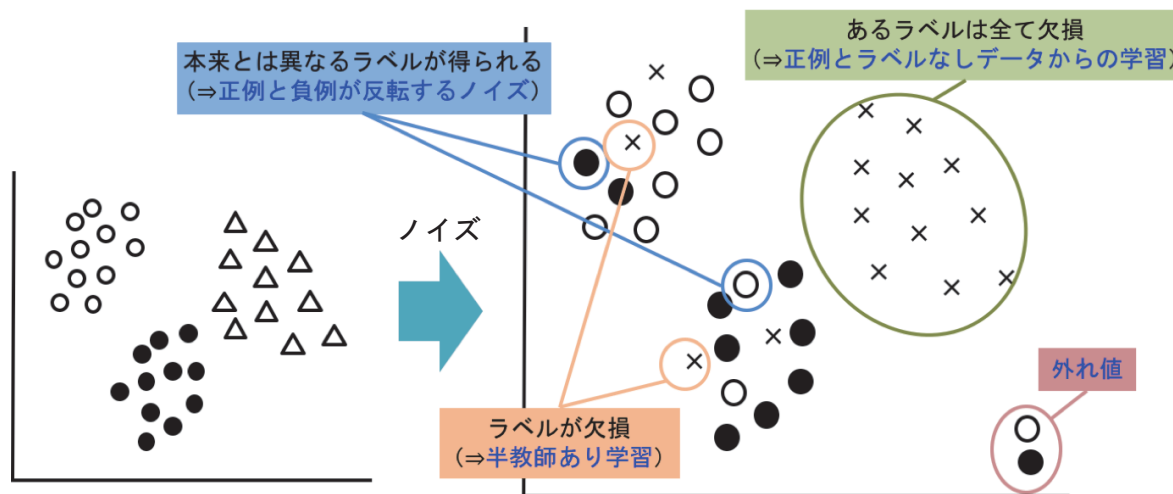
目的 ノイズ分布未知の一般化ラベルノイズモデルに対する分類法の提案

手法 EMアルゴリズムを利用した分類法を提案，数値実験により評価

結果 データ数を増やすとノイズ分布既知の分類法に近づく傾向

一般化ラベルノイズモデル

- 様々なノイズモデルを表現可能
 - 半教師あり学習，正例とラベルなしデータによる学習，
 - 外れ値を含むデータからの学習， e.t.c.



評価実験結果

3パラメータGroup Lassoモデルにおける 経験ベイズ解の解析

吉田 司 (豊橋技科大) 渡辺 一帆 (豊橋技科大)

目的

Group Lasso正則化における経験ベイズ推定解の性質の解明

手法

単純化した可解モデルにおける解析

結果

3パラメータGroup Lassoにおける経験ベイズ推定解の厳密解を導出

回帰モデルの損失関数

$$L(\mathbf{w}) = \frac{1}{2} \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{w}\|^2 + \lambda \|\mathbf{w}\|$$

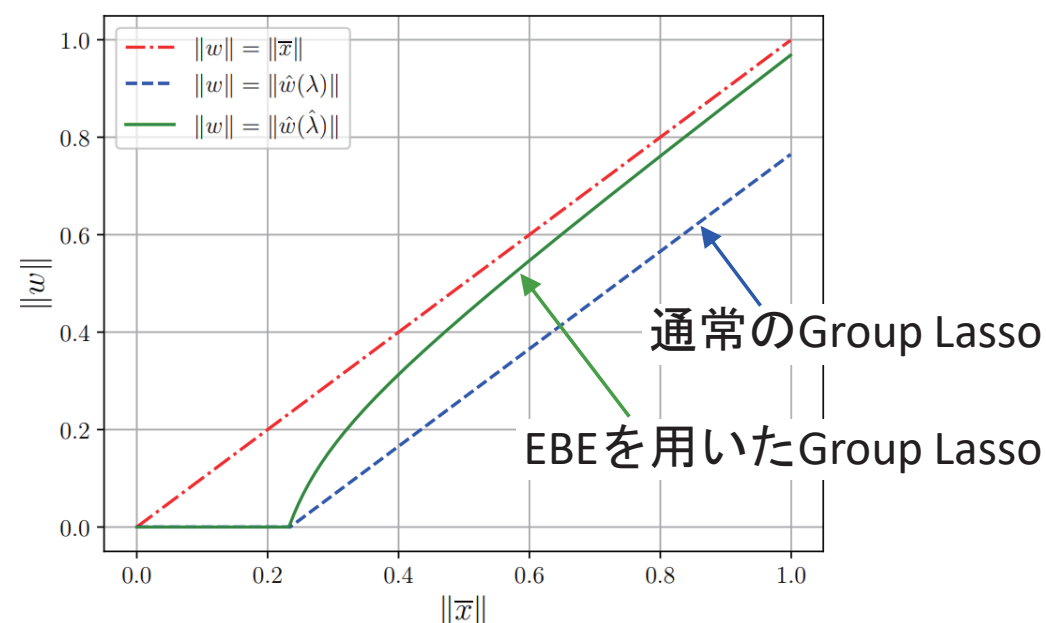
経験ベイズ推定解(EBE)の導出結果

$$\hat{\lambda} = \begin{cases} \infty & \left(\|\bar{\mathbf{x}}\| \leq \sqrt{\frac{3}{n}} \right) \\ \lambda^* & \left(\|\bar{\mathbf{x}}\| > \sqrt{\frac{3}{n}} \right) \end{cases}$$

(ただし, $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$)

さらに λ^* の一意性と,
 λ^* が満たす方程式及び不等式を導出した.

MAP推定解の閾値関数



EBEを用いた場合のMAP解の閾値関数は
Adaptive Lassoの閾値関数と類似性有り

ガウス過程の導関数に基づく極小点の同定のための能動学習

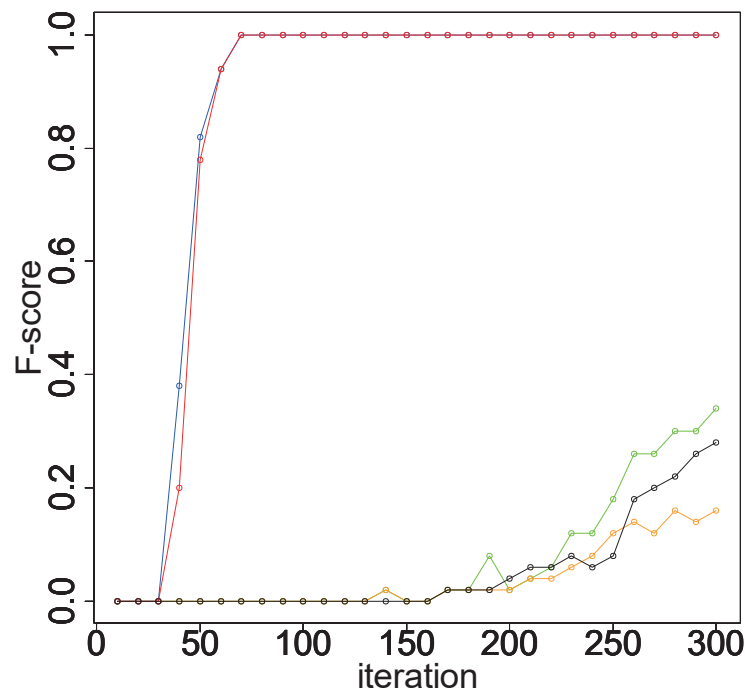
稲津 佑 (理研) 梶田 大輔 (名工大) 豊浦 和明 (京大)

竹内 一郎 (名工大/理研/物質・材料研究機構)

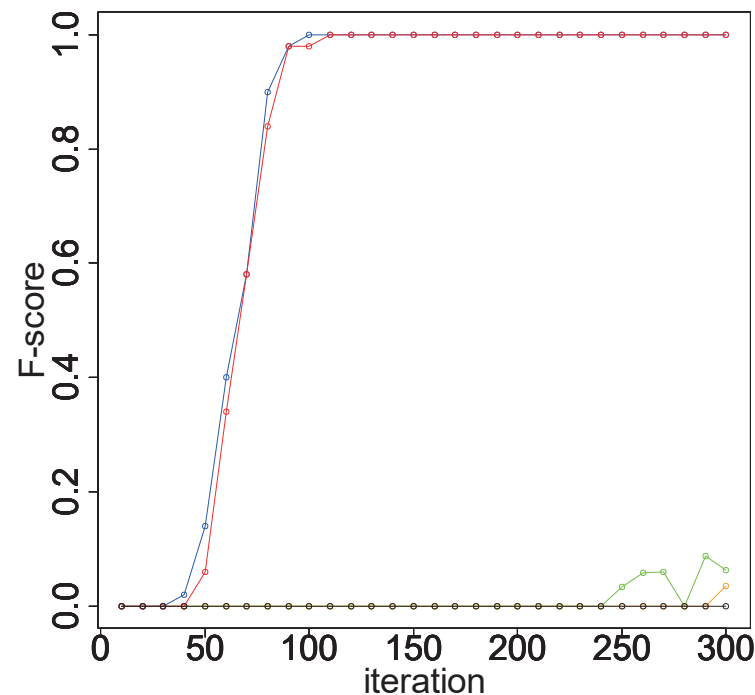
目的 未知関数の極小点を効率的に同定するための能動学習を行う

手法 ガウス過程の微分に基づく勾配の信頼区間を用いたレベルセット推定

結果 全ての点を評価することなく高い精度での極小点の同定に成功



真の関数が2変数関数



真の関数が3変数関数

- 提案法(赤,青)は**少数**の評価回数で高いF値を達成
- 2変数の場合70回, 3変数の場合100回程度でF値が1に到達

Selective Inferenceに基づく スパース線形回帰モデルにおける能動学習

梅津 佑太 (名工大) 竹内 一郎 (名工大, 理研, NIMS)

目的

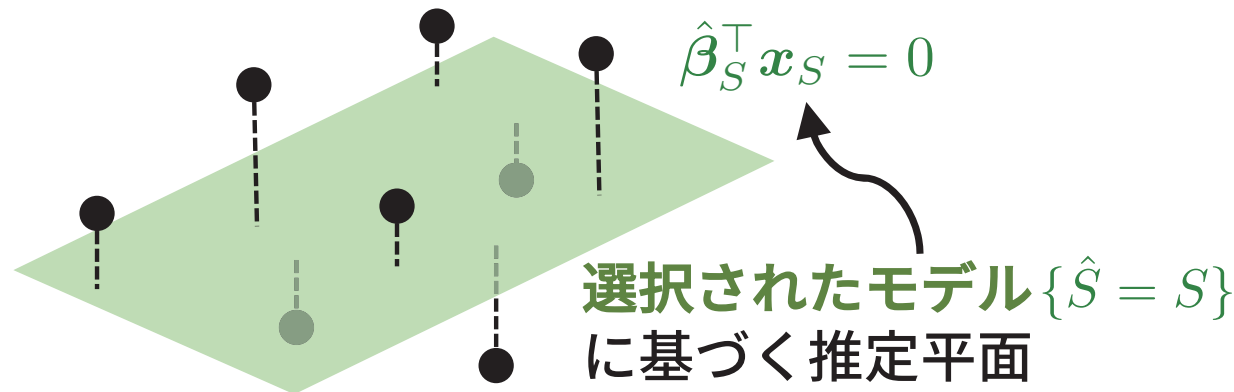
スパース線形回帰モデルの能動学習を
頻度論的に考察

手法

selective inferenceに基づく, MSE最小化
によって入力をサンプリング

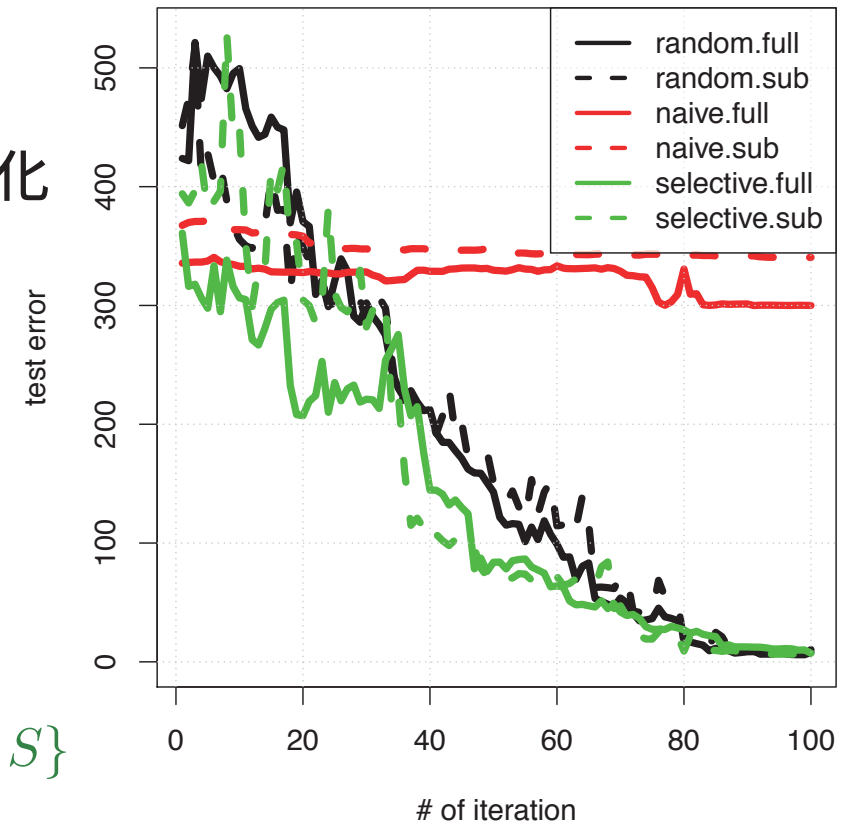
結果

少ない回数でランダムサンプリング
よりも予測誤差を改善



Selective A-optimality

$$\min \mathbb{E}[\|\hat{\beta}_S - \beta_S^*\|^2 \mid \hat{S} = S]$$



緑: 提案手法

黒: ランダムサンプリング

赤: モデル選択の影響を
無視したナイーブな手法

非負値行列因子分解とMDL規準によるマーケットプレイスにおける商品推薦手法

荒野 洋輔 (九大) 三宅 悠介 (GMOペパボ) 川喜田 雅則 (九大) 竹内 純一 (九大)

目的

商品の生産数が少なく一般手法が適用困難なハンドメイドマーケットminneにおける商品推薦

手法

モデル選択にMDL規準を用いた非負値行列因子分解(NMF)によるソフトクラスタリング

結果

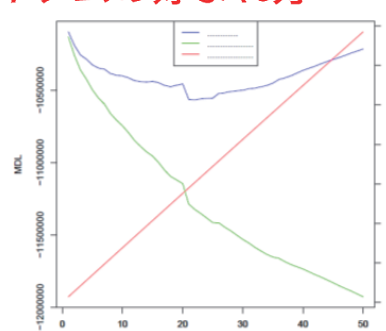
推薦機会率：3割→10割，クリック率：1.25倍

ユーザのお気に入り履歴

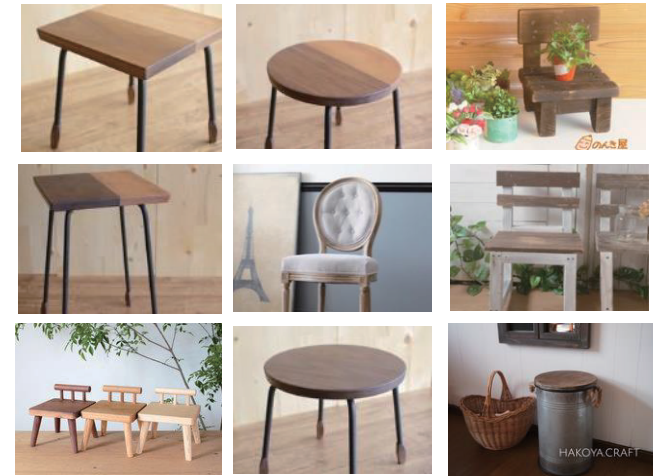
ユーザのお気に入り履歴から
作成したuser-item行列

$$V = \sum_r w_i \times h_i$$

NMFによりuser-item行列を
ユーザの好みとアイテムの好まれ方
に基づいて分解



MDL規準によりクラスタ数を選択



1 クラスタ内の商品写真の例(家具カテゴリ)
適切に商品の特徴が抽出，分類されている

他手法の約1.25倍の
クリック率を実現

統計的決定理論に基づいた 因果効果の推定法に関する一考察

堀井 俊佑 須子 統太

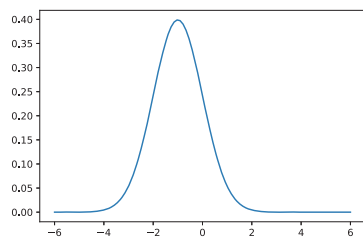
目的 ベイズ基準のもとで最適な介入効果の推定法を導出

手法 介入効果の推定問題を統計的決定理論的に定式化

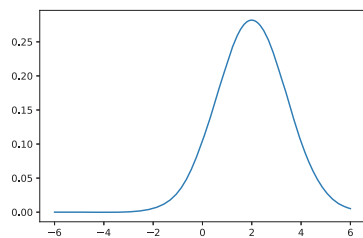
結果 パラメータ・因果ダイアグラムの事後分布で重み付けるのが最適

・ 因果ダイアグラムに基づいた介入効果の一般的な推定手順

1. 因果ダイアグラムの推定 $\Rightarrow$ 2. 条件付き確率分布の推定 $\Rightarrow$ 3. 介入効果の計算

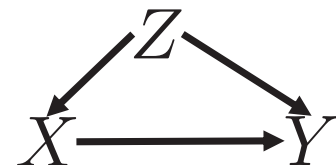


推定介入効果



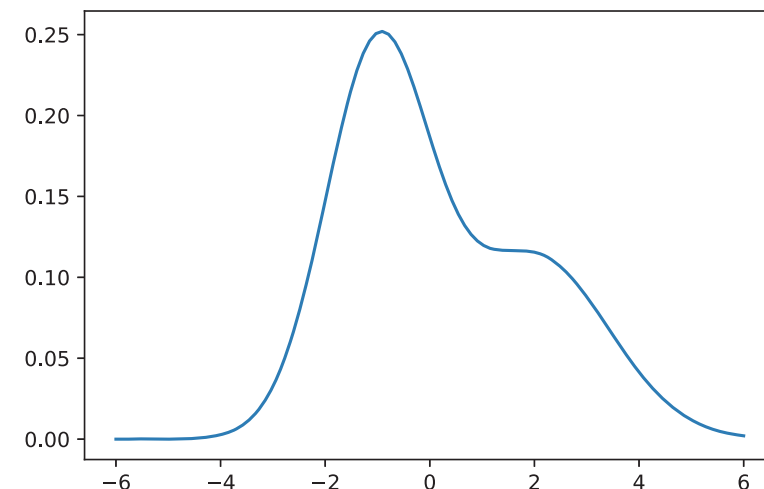
$$X \rightarrow Z \rightarrow Y$$

因果ダイアグラム1



因果ダイアグラム2

ベイズ最適な介入効果推定：



EXAFS(広域 X 線吸収微細構造) の ベイズ的 LARS-OLS を用いたスパースモデリング

五十嵐 康彦 (JSTさきがけ) 岩満一功(熊本大) 岡島敏浩(九州シンクロトロン)
赤井 一郎 (熊本大) 岡田 真人 (東大)

目的 EXAFSデータに対しスパース回帰にて基底選択し、物質内の近接構造を抽出

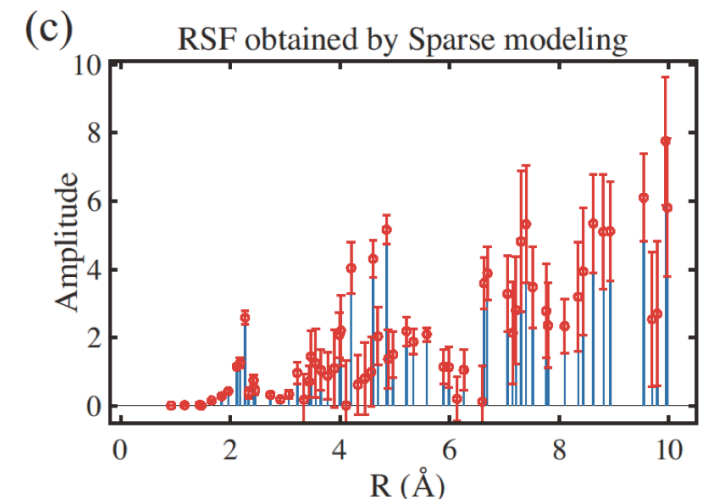
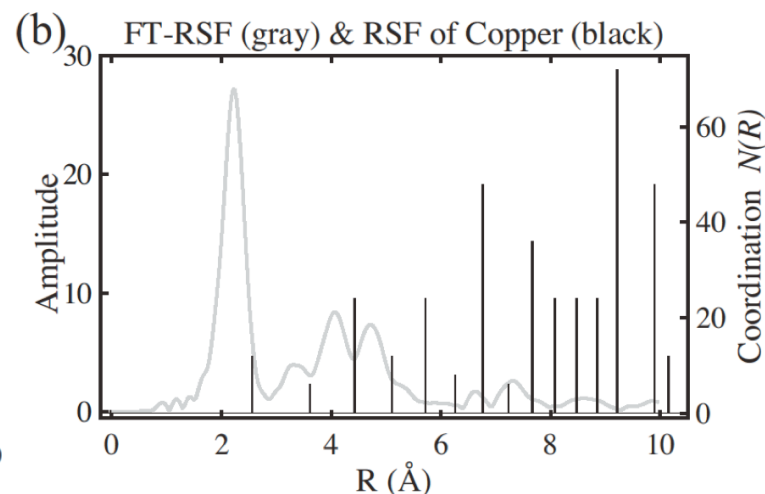
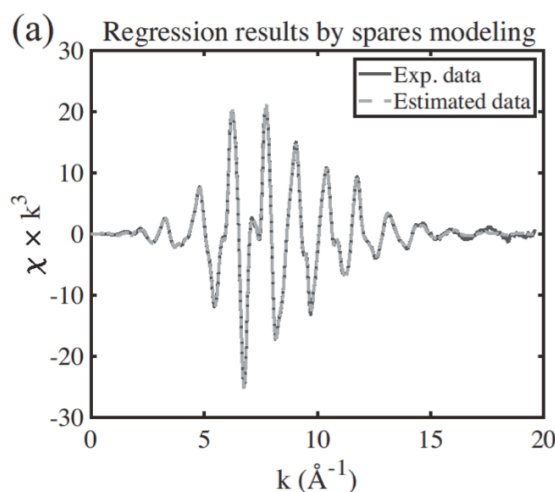
手法 L1正則化パラメータをベイズ的自由エネルギーで評価し、物理モデル抽出

結果 物理モデルと対応する基底が抽出され、その信頼度も分散推定により評価

銅箔に対する
EXAFSデータ
と回帰結果

銅の離散的な近接構造(黒)
とフーリエ変換による
動径構造分布(グレー)

本手法で得られた
動径構造分布(青)と
その分散推定(赤)の結果



ランダム畳込みネットワークとエコーステートネットワークの 組合せによる3D骨格情報からの人間行動認識

トウ シキョウ(東大) 田中 剛平(東大)

目的

Simplify and **accelerate** the training process for **high-dimensional** data classification.

手法

Combine **Untrained Convolutional Neural Networks** and **Echo State Networks**

結果

The numerical experiments show the classification ability of the presented method.

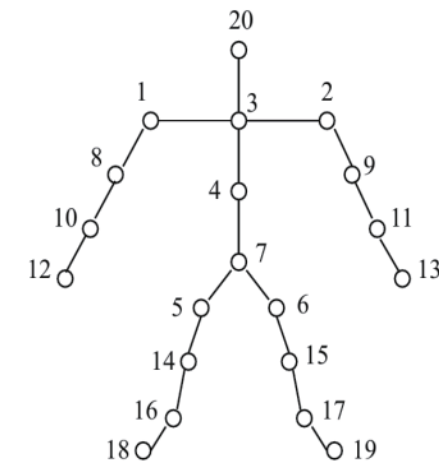
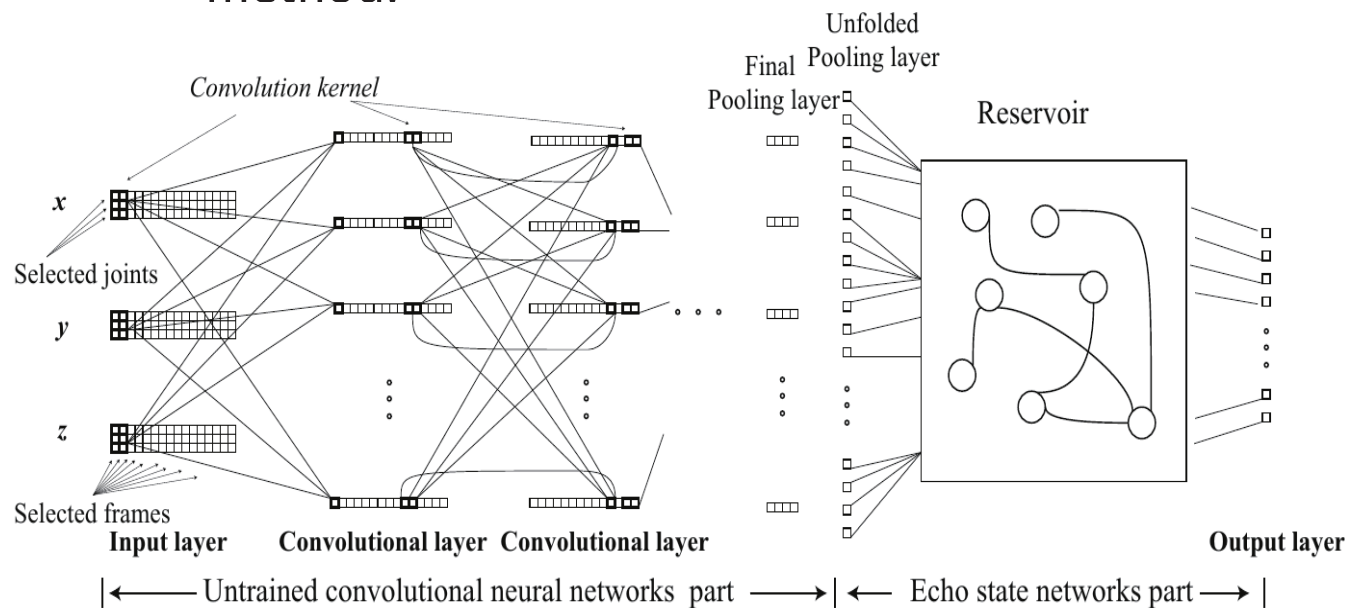


Fig.2. MSR Action 3D Dataset [17].

Fig.1. The computational framework used in this study for human action recognition [10].

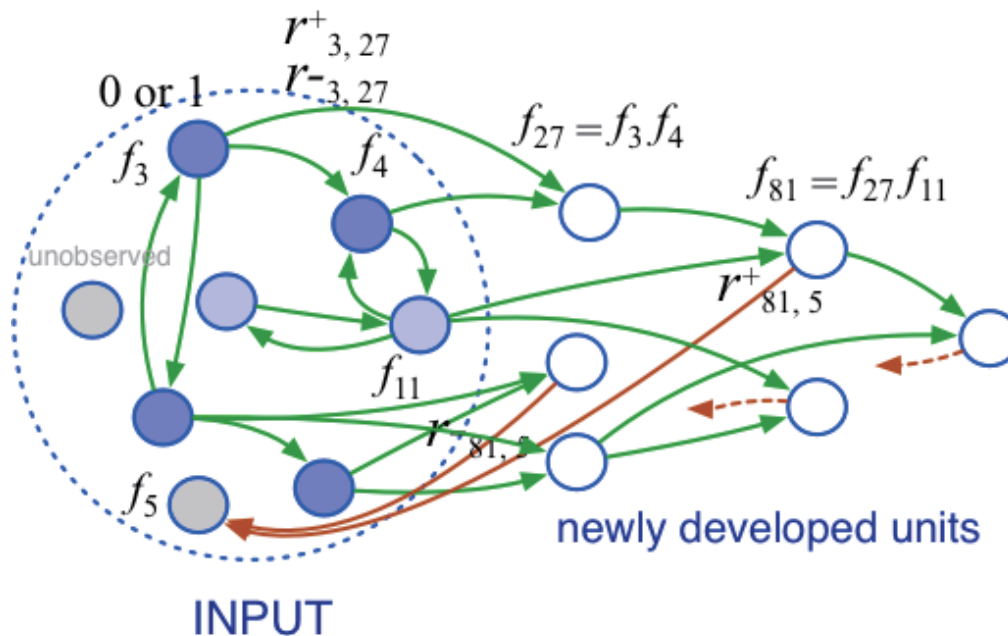
自己テストにより高次素子を追加し学習する自己組織化モデル

伊達 章 (宮崎大) 花井 俊介 (宮崎大)

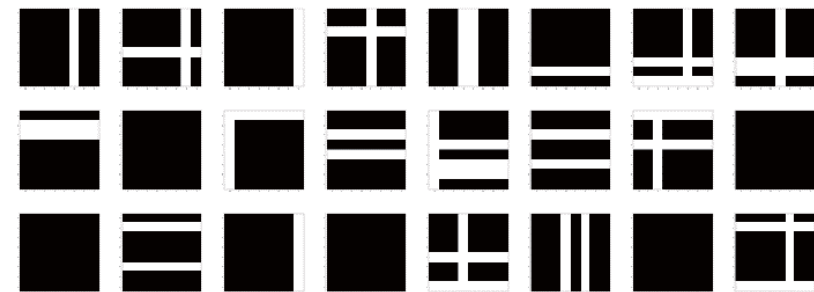
目的 入力信号空間中に潜んでいる構造（空間的規則性）の抽出

手法 1980年代前半に提案された Geman-Davis のモデル を使用

結果 必要とされる高次の素子が追加されていく様子を観察



訓練画像



4次



d



5次

Transfer Learning Among Time Series Data

Hao NIU Kei YONEKAWA Mori KUROKAWA Arei KOBAYASHI (KDDI Research, Inc.)

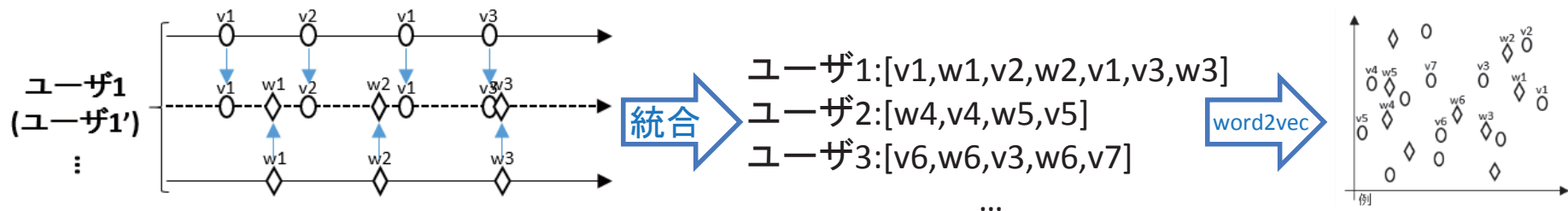
目的 時系列データセット間転移可能な表現を学習

手法 マルチデータセットを統合し、埋め込む

結果 データセット不変な表現なので転移可能

- データセット不変な表現
- データセット間で共通のユーザ情報を用いて時系列データを統合し埋め込む

※共通のユーザ情報：ユーザ i = ユーザ i' 、 $i = 1, 2, \dots, n$



転移効果の評価

- KaggleのRetailrocketデータを2セットに分けて、データセット不変な表現を学習する
- 学習したベクトルを基に一方のデータセットの購買推定モデルを構築し、他方のデータセットで評価する。ROC_AUC>0.5の結果を獲得し、転移可能であることが確認できた。

アンサンブル学習における弱学習器間の相互性指標に基づく 学習法の修正に関する一検討

内海 翔太 亀山 啓輔

目的 相補性と分類精度の関係を用いたアルゴリズムの修正法の検討

手法 分類精度と相補性に基づきデータセットの傾向を分類，改善法を検討

結果 分類に基づく手法の修正法を確認・評価

- アンサンブル学習において
分類精度と弱学習器の相補性に基づき
分類結果を4つにパターン分け

- LL, LH, HL, HH
- 4つのパターンそれぞれを考察し，
学習法の修正策を検討
- 学習法の変動によりパターンごとの性質が考察通りであることを確認
- LHのパターンを持つデータセットの分類精度の改善を確認

相補性 \ 分類精度	低	高
	LL	LH
低		
高	HL	HH

ドメイン敵対的訓練を用いた 半教師あり異種転移可能表現学習におけるmixupの効果

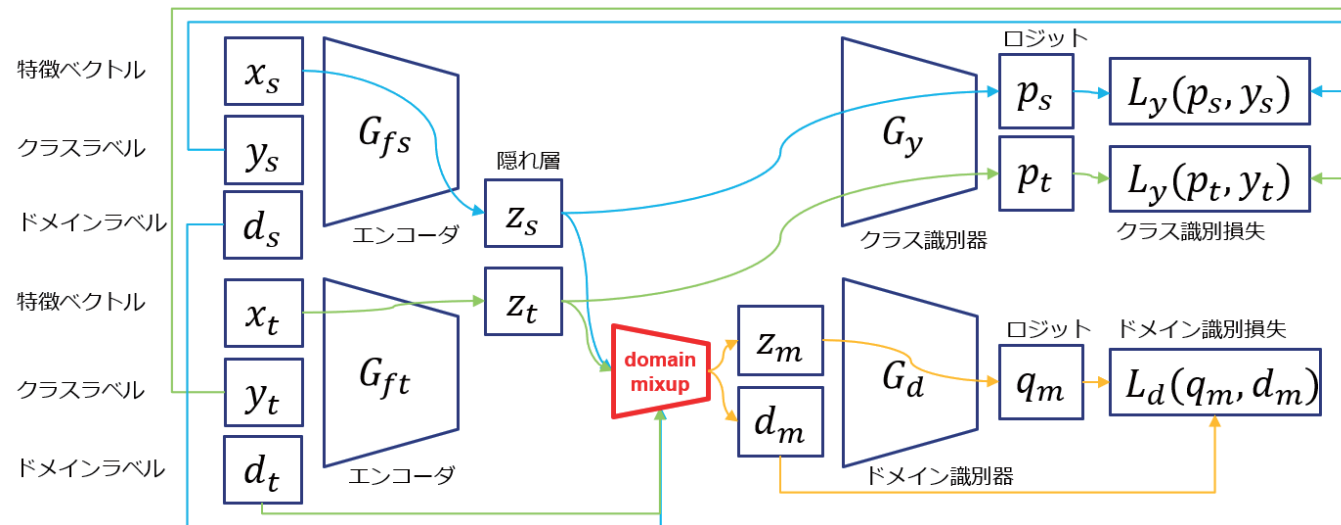
米川 慧 牛 コウ 黒川 茂莉 小林 亜令 (KDDI総合研究所)

目的 異種転移学習における特徴表現のドメイン不変性向上

手法 ドメイン敵対的訓練のドメイン識別器にデータ強化手法mixupを適用

結果 目標ドメインのラベルあり事例が少ないケースで精度向上を確認

方法



評価

- データ : Breast Cancer Wisconsin (乳癌に関する患者の検体細胞の統計量と診断結果)
 - Original: 医師の判断による統計量→元ドメインとする
 - Diagnostic: 画像処理による統計量→目標ドメインとする
- 目標ドメインのラベルありサンプル率を変えて正答率を比較→**最大+4%精度向上**

DNNの捉えた多様体構造からの保存量推定

～ 多様体の対称性抽出とネーターの定理を介した保存量推定 ～

本武陽一（東大）

目的

力学系時系列データからの系が持つ保存量の推定

手法

力学系の時系列データを学習したDeep Auto Encoderによる対称性抽出

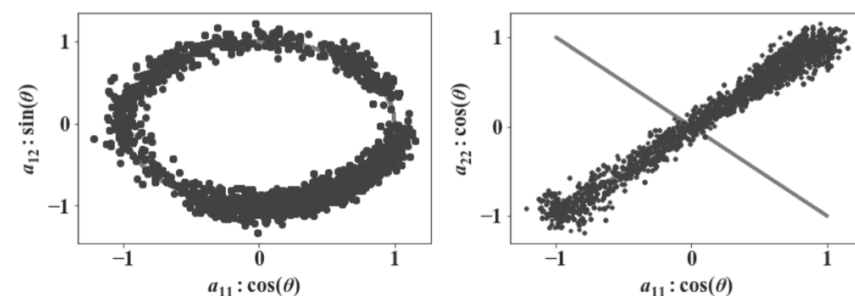
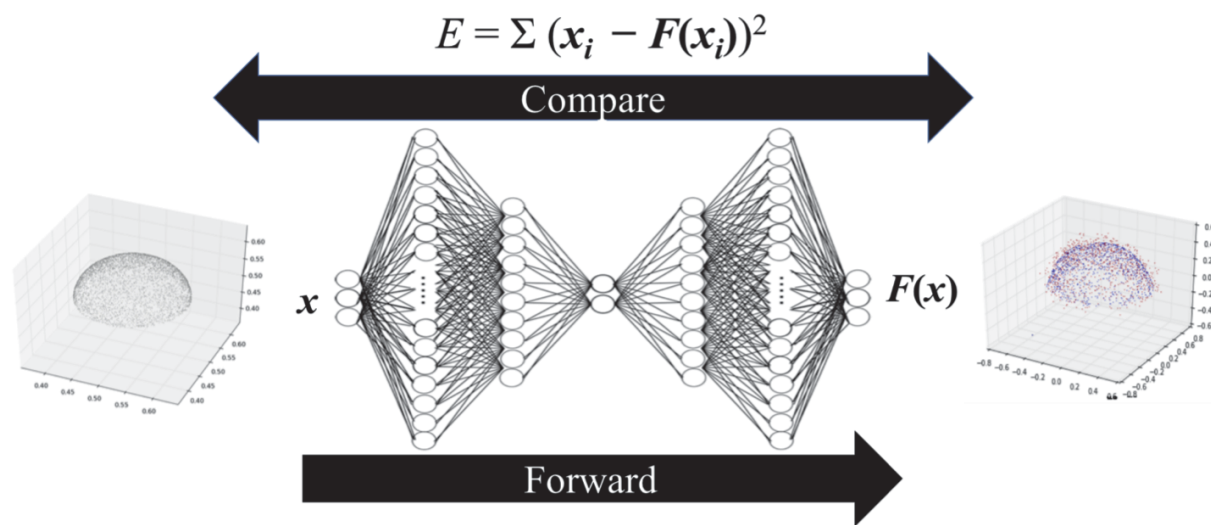
結果

中心力ポテンシャルで運動する系の角運動量保存則の抽出が実現された

- 提案手法：Deep Auto Encoderを用いた対称性抽出

- 結果：

系を変えない座標変換 $\mathbf{A}$ の推定結果



$$\mathbf{A} = \begin{pmatrix} \cos(\theta_z) & \sin(\theta_z) \\ -\sin(\theta_z) & \cos(\theta_z) \end{pmatrix}$$

- ⇒ 入力空間中の多様体領域を引き戻し写像（DAE）の同相性から抽出（ E が小さくなる=多様体上）
- ⇒ 角運動量保存則に対応する回転対称性が抽出された