

T2-01

Analysis of Dropout in Online Learning

Kazuyuki Hara Nihon Univ.

Dropout is effective in online learning ?

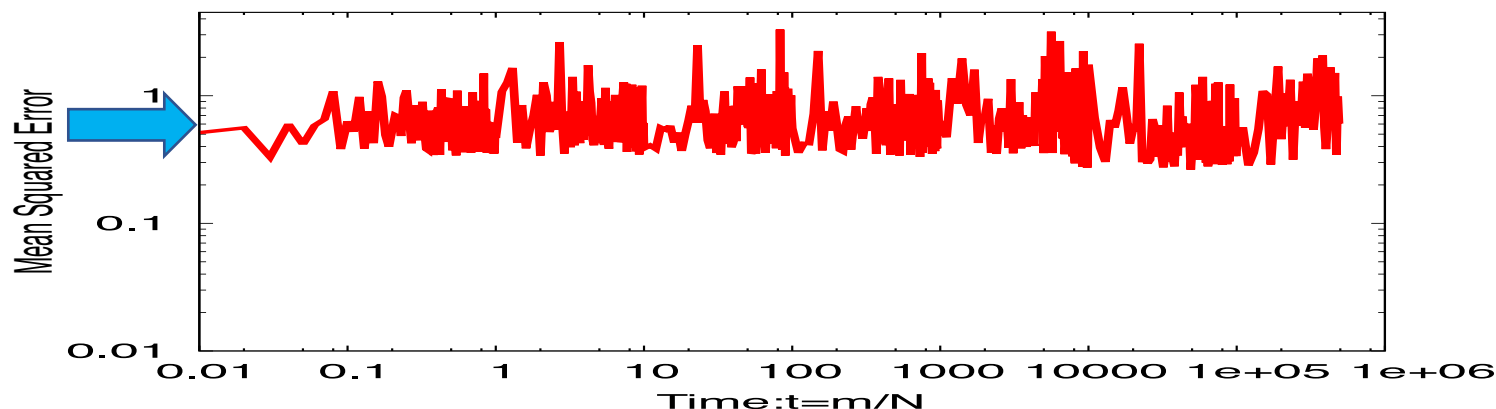


Apply Dropout to the three layer network
and watch
effect of Dropout for redundant network

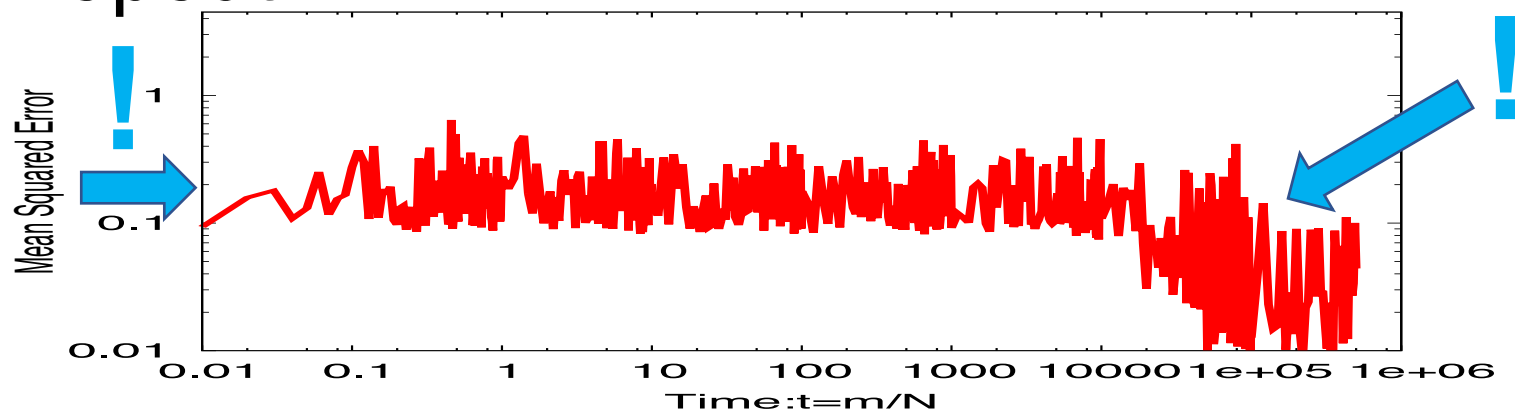


Dropout have some effect of reducing
redundancy of the network

Error is not decreased by redundant network without Dropout



Error is decreased by redundant network with Dropout



More about this, please visit my poster [T2-01](#)

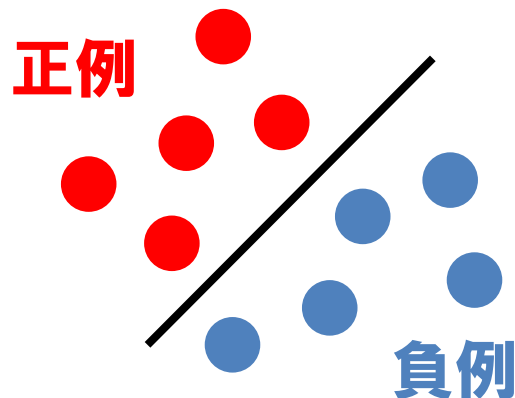
正例度データからの二値分類

石田 隆 (1,2,3) ニュー ガン (2,3) 杉山 将 (3,2)

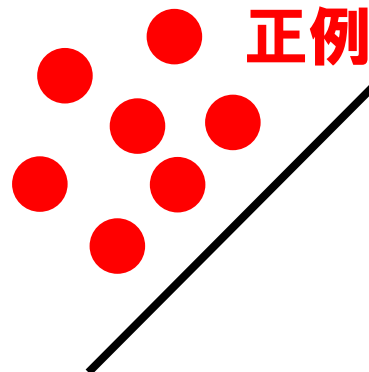
(1) 三井住友アセット
(2) 東京大学
(3) 理化学研究所

正例は簡単に集められるが、負例は集められない！
そんな時、どうする？

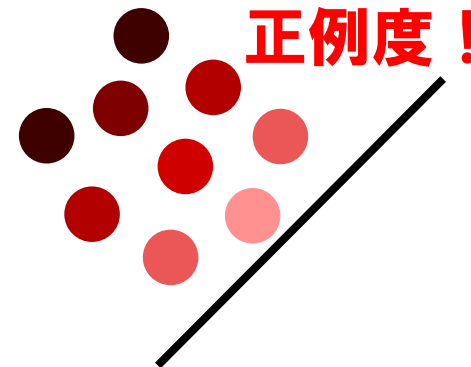
通常の分類



解きたい問題



提案法



正例度の付いた**正例**から分類器を学習しよう！

経験リスク最小化で実現、最適な $1/\sqrt{n}$ 収束を達成

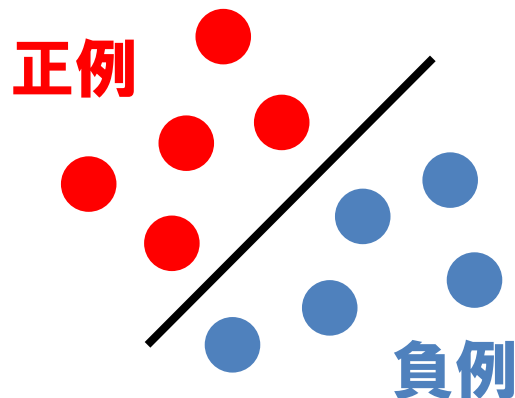
正例度データからの二値分類

石田 隆 (1,2,3) ニュー ガン (2,3) 杉山 将 (3,2)

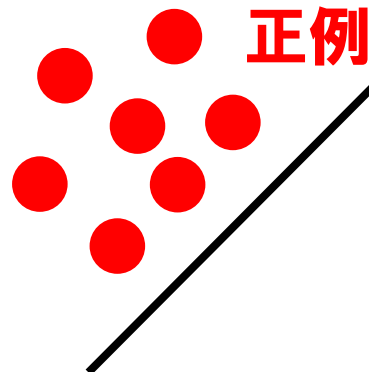
(1) 三井住友アセット
(2) 東京大学
(3) 理化学研究所

正例は簡単に集められるが、負例は集められない！
そんな時、どうする？

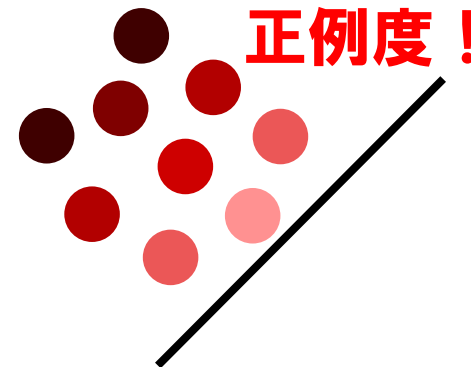
通常の分類



解きたい問題



提案法



正例度の付いた**正例**から分類器を学習しよう！

経験リスク最小化で実現、最適な $1/\sqrt{n}$ 収束を達成

CNN の特徴マップを用いた 回転角推定の連続実行による回転画像認識 (1/2)

片山 錦・山根 智 (金沢大学)

◆ 研究目的

CNNを使った一般物体認識において2次元の回転多様性を吸収する

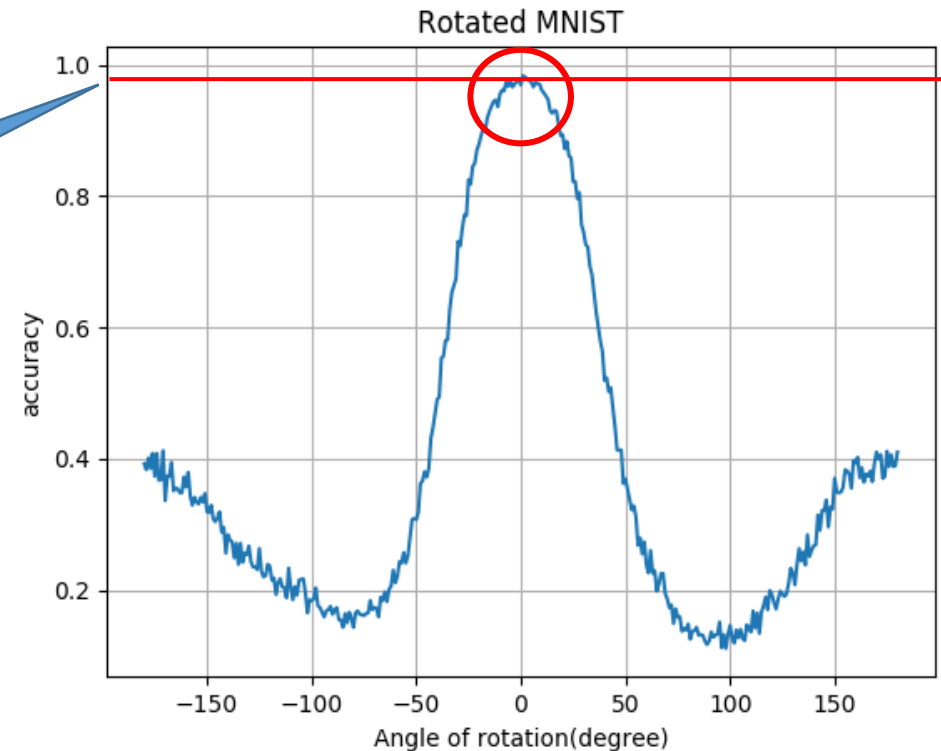
◆ 研究概要

元のCNNを変更しない角度推定ネットワークの追加したモデルの提案

手書き数字のデータセット
MNISTでの一般的なCNNの
クラス分類

画像の回転のない識別は
高性能

画像の回転がある識別は
低性能



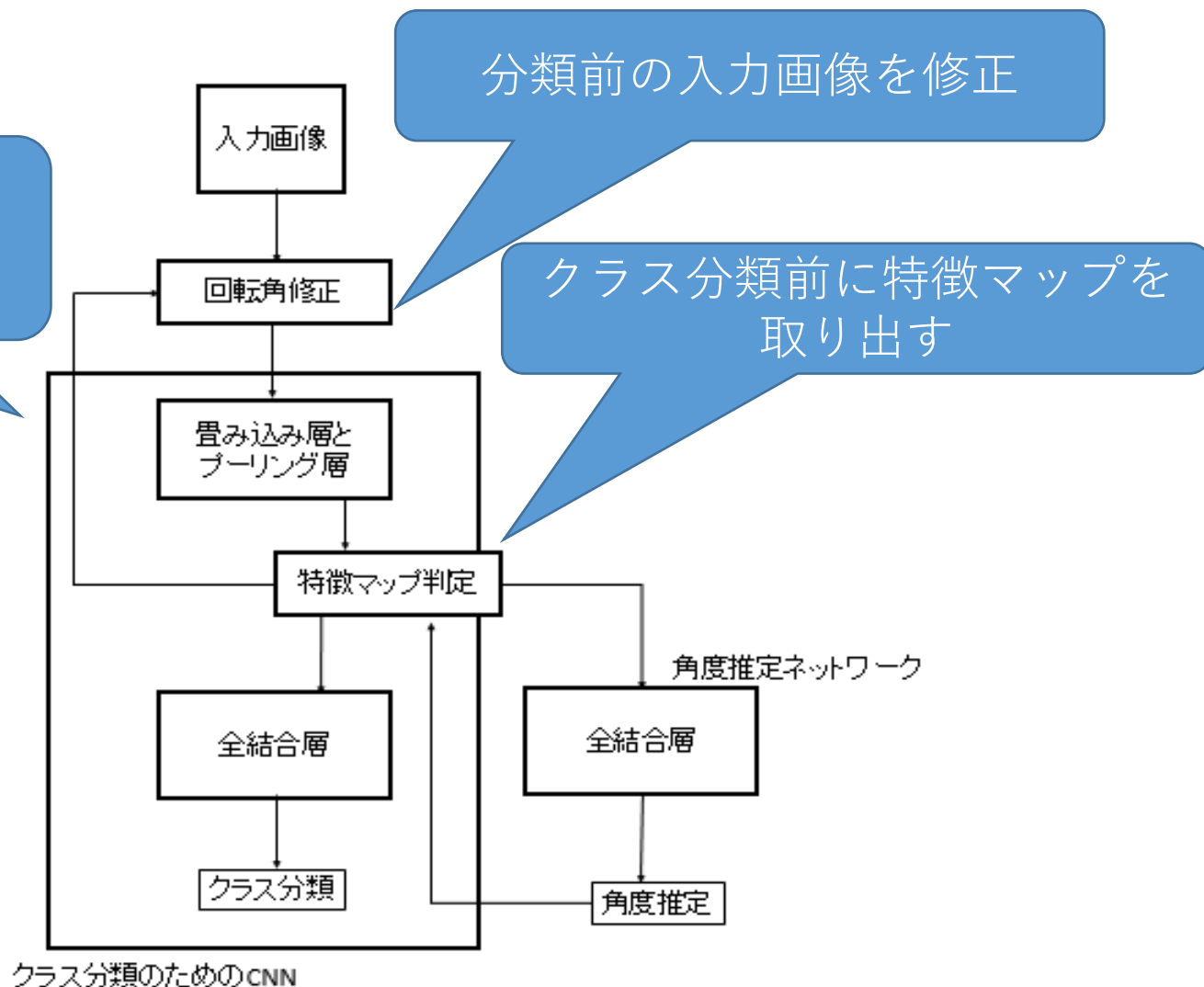
CNN の特徴マップを用いた 回転角推定の連続実行による回転画像認識 (2/2)

片山 錦・山根 智 (金沢大学)

◆ 提案モデル

元々のCNNは変更しない
画像の回転無し
→高精度な認識

回転角の推定・修正を
複数回実行すること
による精度上昇



マージン最大化距離学習におけるセーフスクリーニング

吉田 知貴[†], 竹内 一郎^{†,††,†††}, 烏山 昌幸^{†,††,†††}

学生発表 T2-04

[†]名古屋工業大学, ^{††}物質・材料研究機構, ^{†††}理化学研究所, ^{††††}科学技術振興機構

マージン最大化距離学習の概要と問題点

- 距離学習は、分類しやすくなるようなマハラノビス距離を学習する
- 最適化問題を解くには非常に多くのサンプルの3つ組 (triplet) を考えなければならない
- 小さいデータセットでも triplet は膨大になり、学習に時間がかかる
- 実際に必要な triplet は極わずか

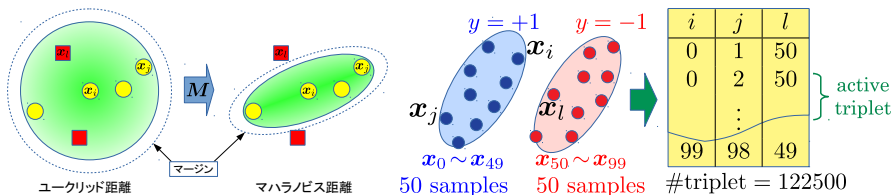


Figure: マージン最大化距離学習

Figure: 膨大な triplet

マージン最大化距離学習におけるセーフスクリーニング

提案法

- 距離学習の制約 (半正定値制約) を考慮し、最適解を包含する超球を導出
- その超球を利用して、最適解に影響を与えない triplet を排除 (Screening) して高速化を図る

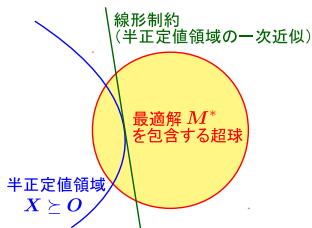


Figure: 超球のイメージ図

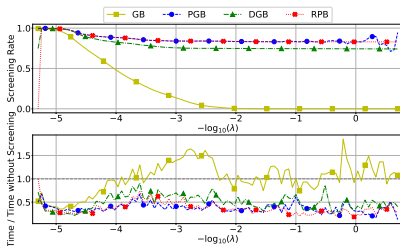


Figure: 導出した 4 つの超球の実験結果

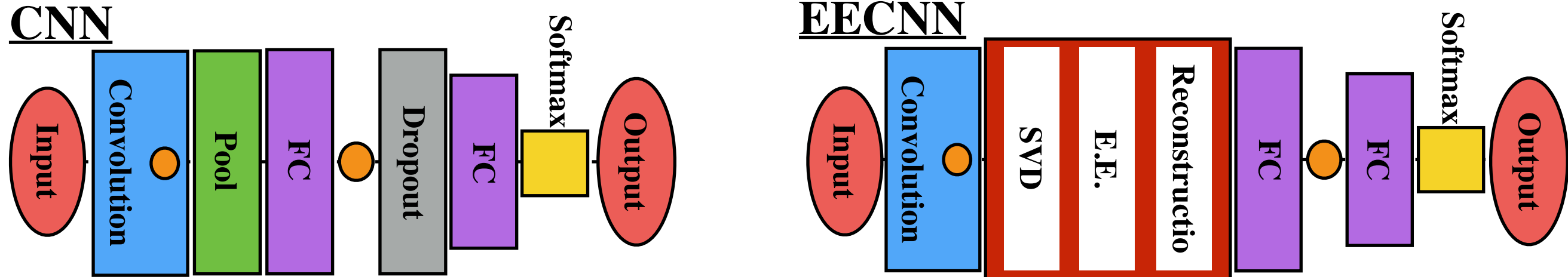
最適解の存在範囲を “超球”, “超球+半正定値制約”, “超球+線形制約” としたときのそれぞれの **Screening Rule** について実験・検証する

T2-05：エンタングルメント・エントロピーを用いた 畳み込みニューラルネットワーク

学生発表

(Convolutional Neural Network with Entanglement Entropy)

○Shu Eguchi, Masaru Tanaka(Fukuoka Univ.)

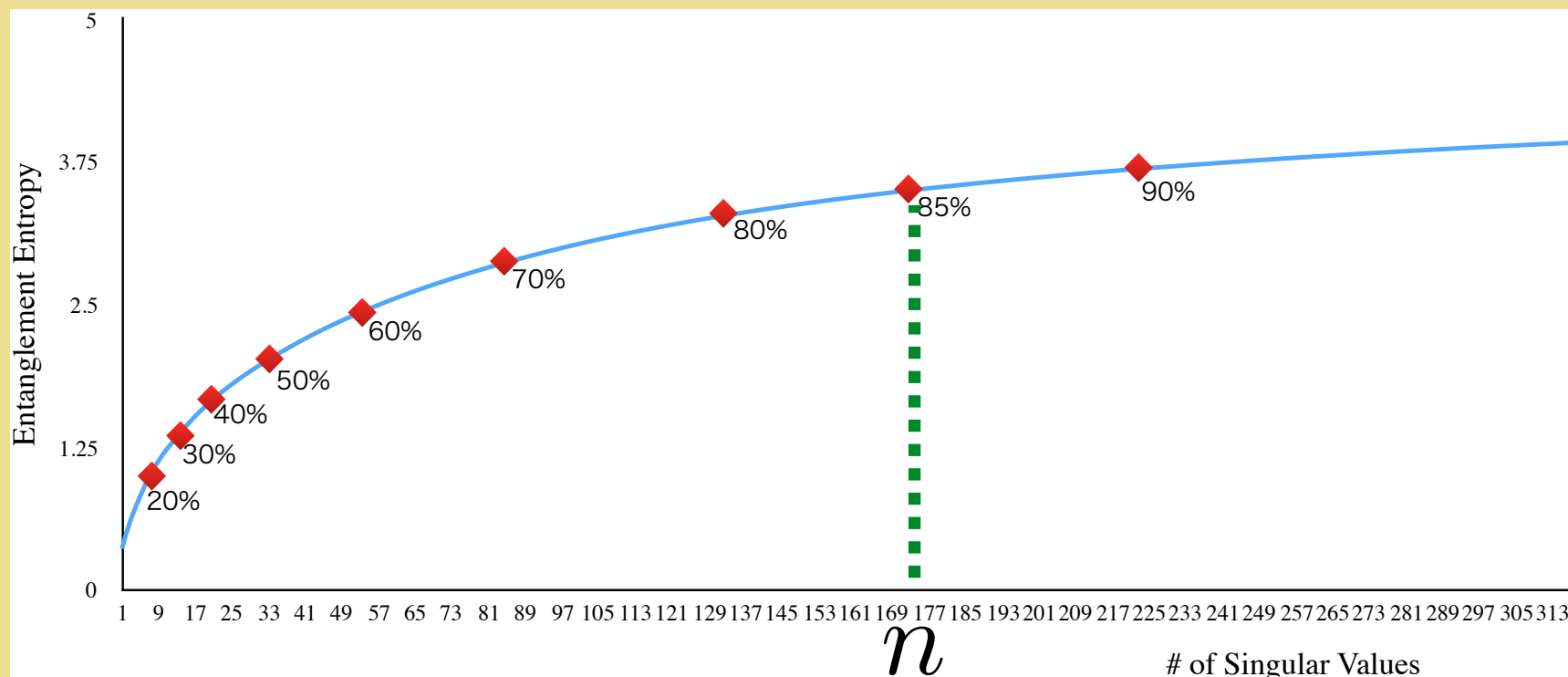


Input



$$= USV$$

Entanglment Entropy



Output

$$US_n V =$$

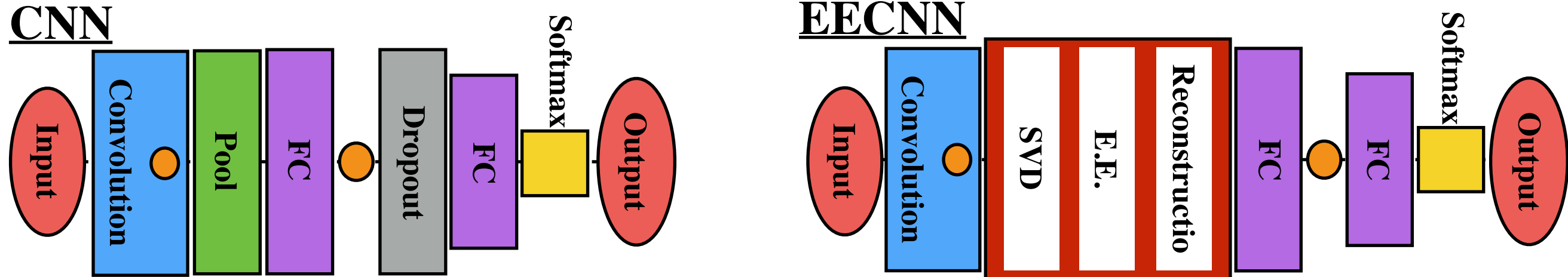


T2-05：エンタングルメント・エントロピーを用いた 畳み込みニューラルネットワーク

学生発表

(Convolutional Neural Network with Entanglement Entropy)

○Shu Eguchi, Masaru Tanaka(Fukuoka Univ.)

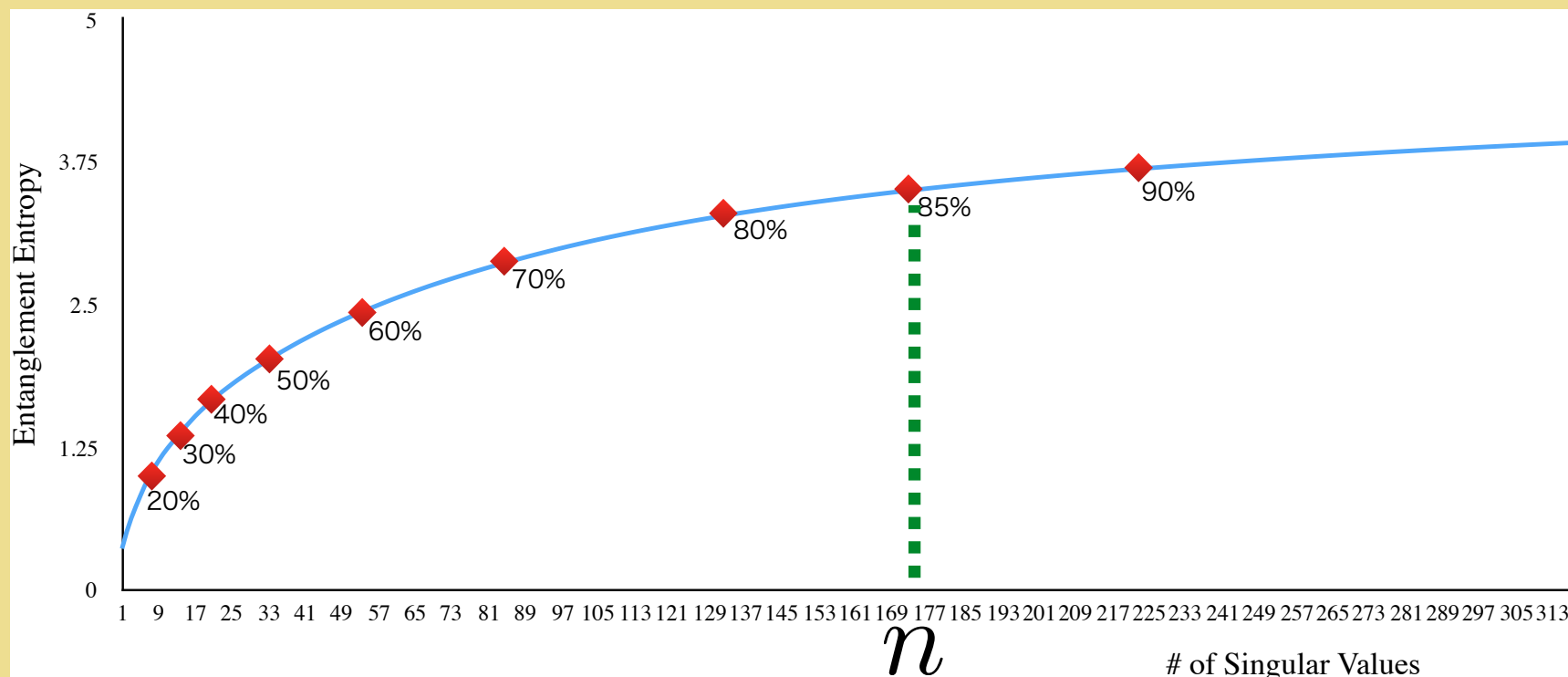


Input



$$= USV$$

Entanglment Entropy



Output

$$US_nV =$$



T2-06: Multi Task Learning with Positive and Unlabeled Data and Its Application to Mental State Prediction

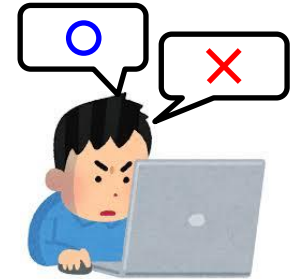
Hiroataka Kaji, Hayato Yamaguchi (Toyota), and Masashi Sugiyama (RIKEN/UTokyo)

Motivation:

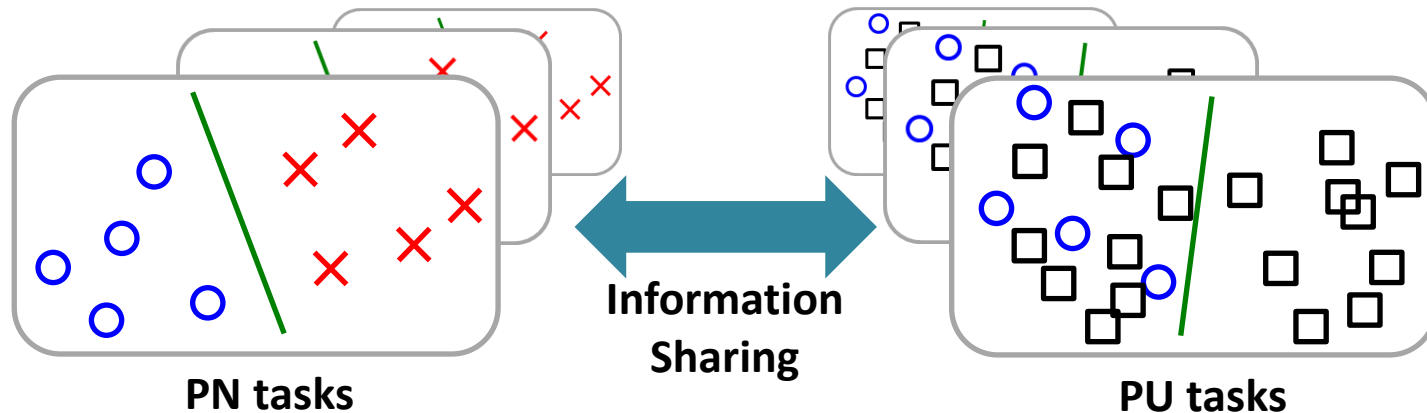
- Learning from a small number of labeled samples

Proposed Method (PU-MTL):

- Handling **Positive-Negative** tasks and **Positive-Unlabeled** tasks in multi-task learning framework

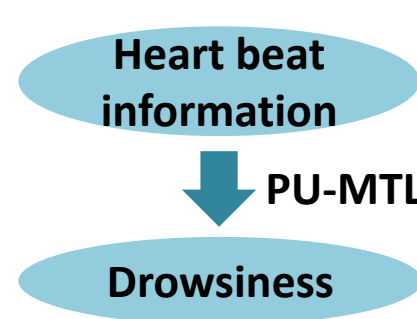


Annotation -> High costs



Application:

- Real-world drowsiness prediction with driving simulator



T2-06: Multi Task Learning with Positive and Unlabeled Data and Its Application to Mental State Prediction

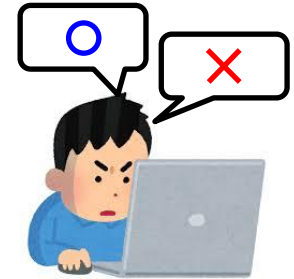
Hiroataka Kaji, Hayato Yamaguchi (Toyota), and Masashi Sugiyama (RIKEN/UTokyo)

Motivation:

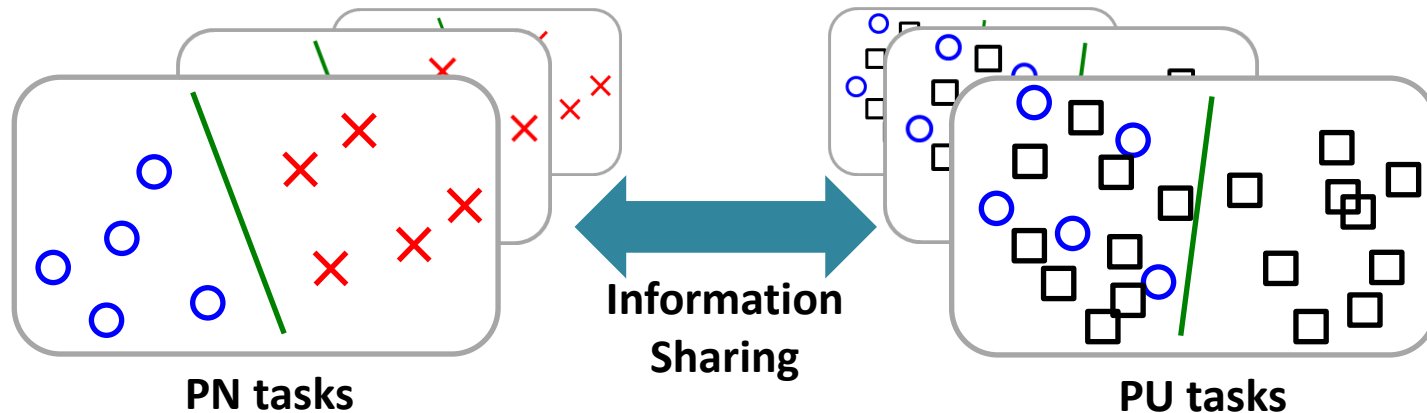
- Learning from a small number of labeled samples

Proposed Method (PU-MTL):

- Handling **Positive-Negative** tasks and **Positive-Unlabeled** tasks in multi-task learning framework

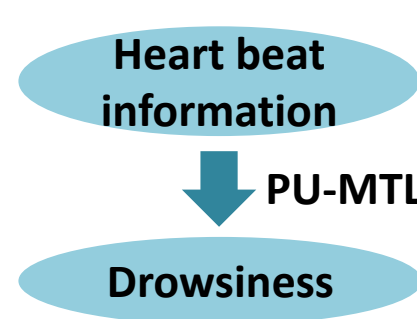


Annotation -> High costs



Application:

- Real-world drowsiness prediction with driving simulator

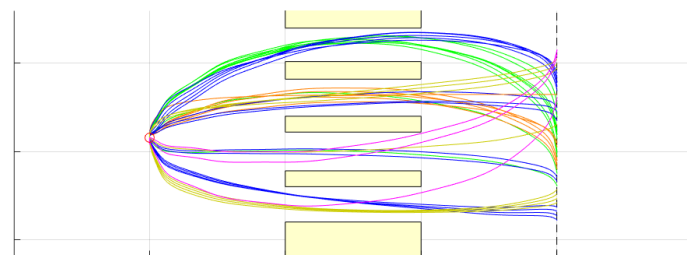
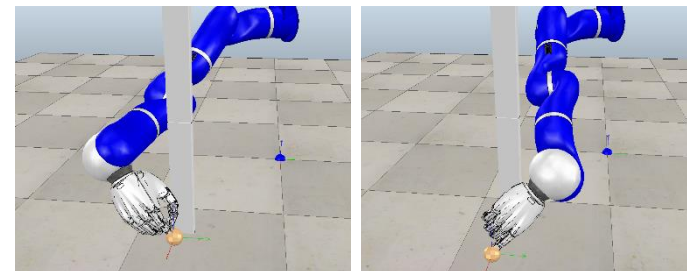
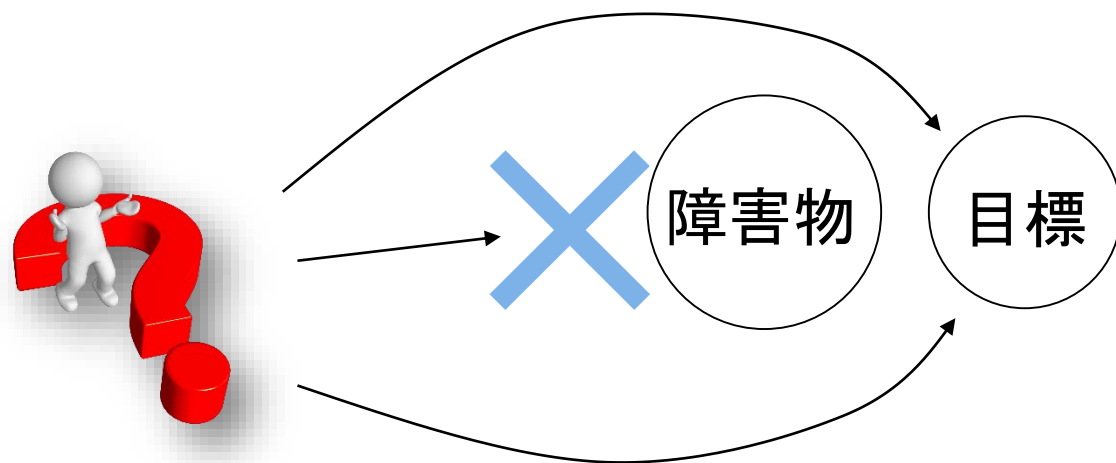


報酬重み付き密度推定に基づく階層強化学習

長 隆之^{†‡} 杉山将^{‡†}

[†] 東京大学
[‡] 理化学研究所

背景: 現実世界では正しい方策が複数存在する
中間をとるような方策はうまくいかない



複数の方策を学習し, 状況に応じて適切な方策を使い分ける

➡ 階層強化学習

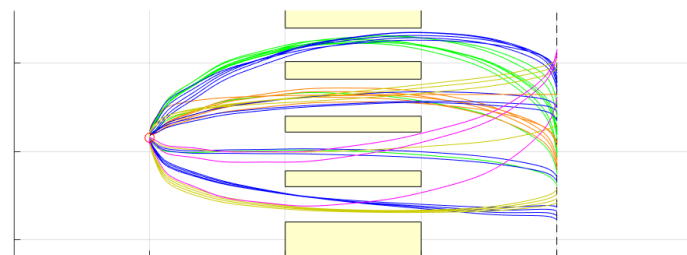
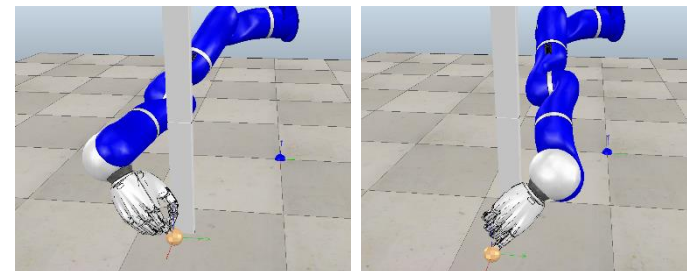
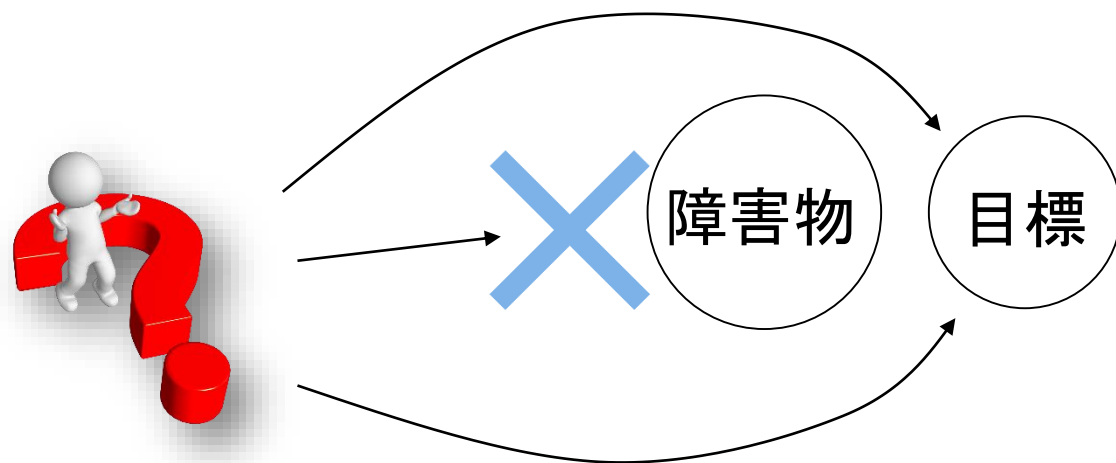
提案手法: 強化学習を報酬重み付き密度推定として捉え, 最適な方策の構造を学習する

報酬重み付き密度推定に基づく階層強化学習

長 隆之^{†‡} 杉山将^{‡†}

[†] 東京大学
[‡] 理化学研究所

背景: 現実世界では正しい方策が複数存在する
中間をとるような方策はうまくいかない



複数の方策を学習し, 状況に応じて適切な方策を使い分ける

➡ 階層強化学習

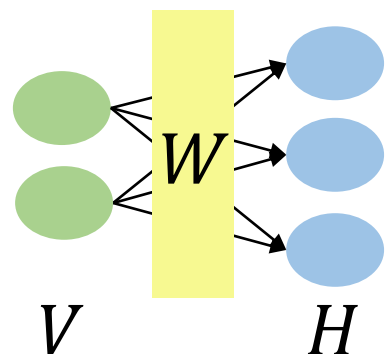
提案手法: 強化学習を報酬重み付き密度推定として捉え, 最適な方策の構造を学習する

PCD法に基づいた畳み込み制約付きボルツマンマシンの学習法の改良

石 涼介¹, 須田 玲輝², 竹田 晃人³ (茨城大院理工^{1,3}, 茨城大工²)

目的: PCD法における学習パラメータの最適値の調査

畳み込み制約付きボルツマンマシン



入力層と出力層により構成される確率的生成モデル

V は入力層, H は出力層を表し, V と H の間に重み W が存在

学習則

PCD法における W の勾配 ΔW は以下のように定義

$$\Delta W = V * H - \hat{V} * \hat{H}$$

$$W^{(t+1)} \leftarrow W^{(t)} + \alpha(\Delta W^{(t)} + \beta \Delta W^{(t-1)} + \gamma |W^{(t)}|)$$

*: 畳み込み演算

$\hat{V}, \hat{H}$: 復元した V, H

α : 学習率

β : モーメント係数

γ : 荷重減衰係数

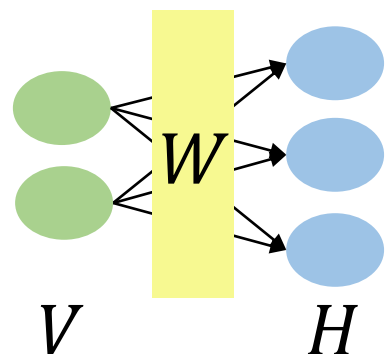
数値実験により α, β, γ の最適値を考察し, 一般的な学習法であるCD法との比較を行う

PCD法に基づいた畳み込み制約付きボルツマンマシンの学習法の改良

石 涼介¹, 須田 玲輝², 竹田 晃人³ (茨城大院理工^{1,3}, 茨城大工²)

目的: PCD法における学習パラメータの最適値の調査

畳み込み制約付きボルツマンマシン



入力層と出力層により構成される確率的生成モデル

V は入力層, H は出力層を表し, V と H の間に重み W が存在

学習則

PCD法における W の勾配 ΔW は以下のように定義

$$\Delta W = V * H - \hat{V} * \hat{H}$$

$$W^{(t+1)} \leftarrow W^{(t)} + \alpha(\Delta W^{(t)} + \beta \Delta W^{(t-1)} + \gamma |W^{(t)}|)$$

*: 畳み込み演算

$\hat{V}, \hat{H}$: 復元した V, H

α : 学習率

β : モーメント係数

γ : 荷重減衰係数

数値実験により α, β, γ の最適値を考察し, 一般的な学習法であるCD法との比較を行う

T2-09 カウントデータに関する多次元ヒストグラムのビン幅最適化

武藤健介(東大)・坂本浩隆(東大/学振)・松浦慧介(東大/学振)

有馬 孝尚(東大/CEMS)・岡田真人(東大/NIMS)

学生優秀プレゼンテーション賞対象

背景

実験データの前処理として、スライスして切り出した低次元データに対し、研究者が経験的にビン幅を設定してヒストグラムの作成を行っている。

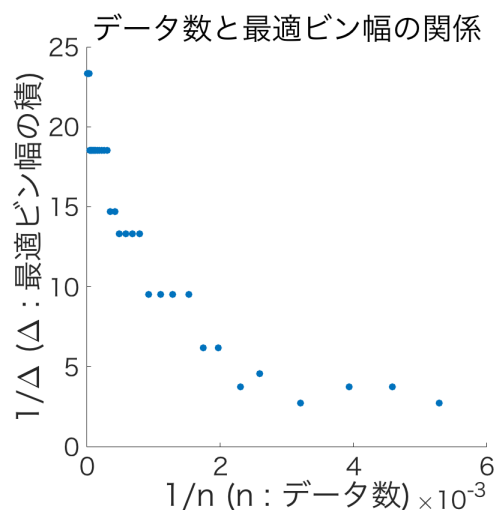
方法

多次元のカウントデータに対して、ビン幅最適化を自動化するアルゴリズムの開発を行った

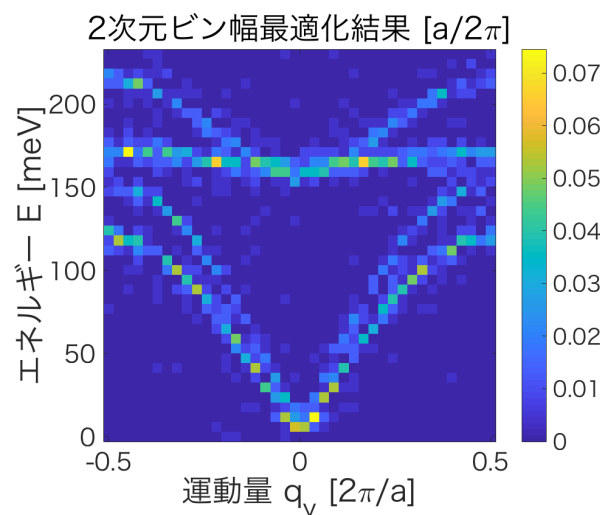
結果

高次元データに対しては低次元のスライスではなく、高次元データ全体に対してビン幅最適化をすべきであることを示した。またデータ数の増加に伴い、小さなビン幅が選択されることが分かった。

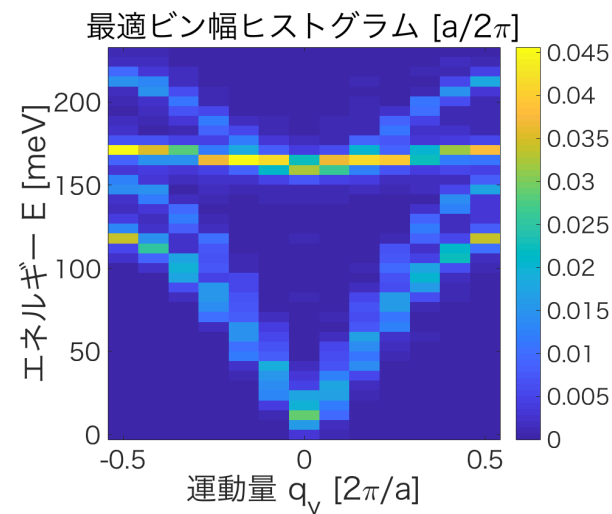
(A) データ数と最適ビン幅



(B) 従来手法



(C) 提案手法



T2-09 カウントデータに関する多次元ヒストグラムのビン幅最適化

武藤健介(東大)・坂本浩隆(東大/学振)・松浦慧介(東大/学振)

有馬 孝尚(東大/CEMS)・岡田真人(東大/NIMS)

学生優秀プレゼンテーション賞対象

背景

実験データの前処理として、スライスして切り出した低次元データに対し、研究者が経験的にビン幅を設定してヒストグラムの作成を行っている。

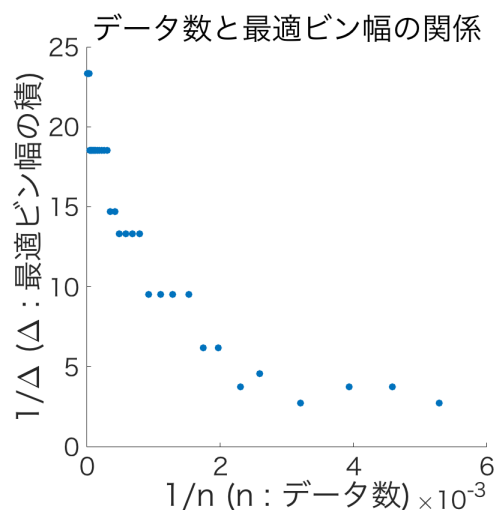
方法

多次元のカウントデータに対して、ビン幅最適化を自動化するアルゴリズムの開発を行った

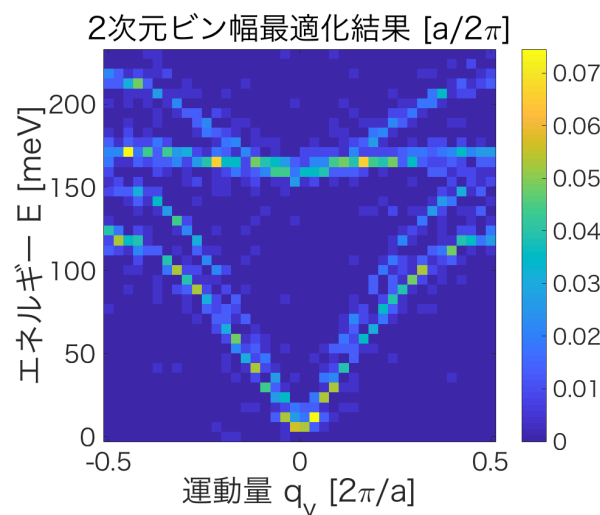
結果

高次元データに対しては低次元のスライスではなく、高次元データ全体に対してビン幅最適化をすべきであることを示した。またデータ数の増加に伴い、小さなビン幅が選択されることが分かった。

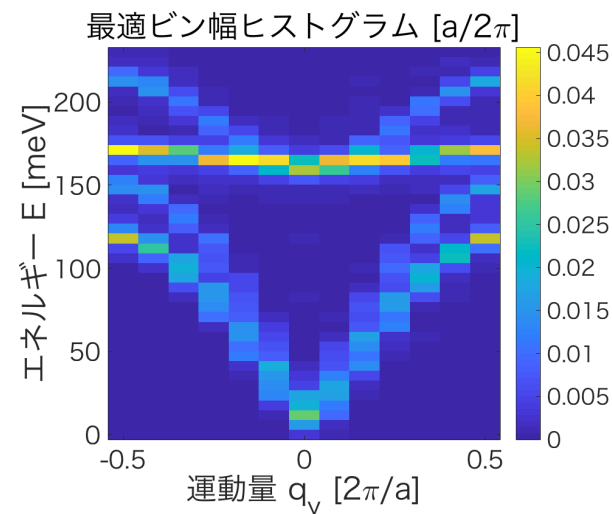
(A) データ数と最適ビン幅



(B) 従来手法



(C) 提案手法



T2-10 べき乗則をもとにした津波高予測のための非線形パラメトリックモデル

吉川真史 (東大), 五十嵐康彦 (JST), 村田 伸 (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)

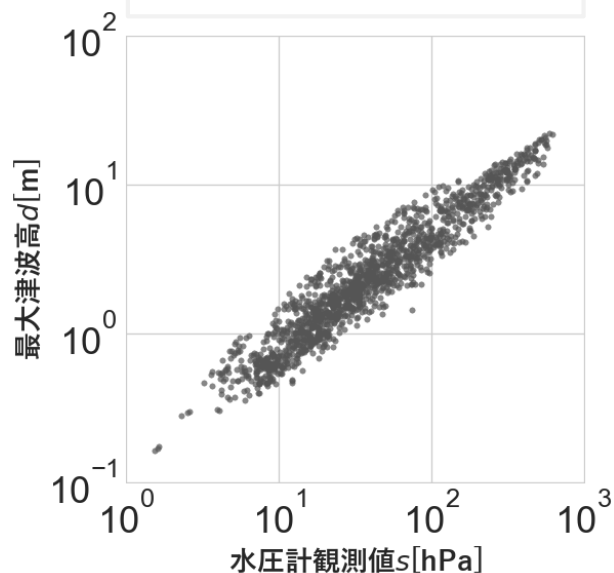
目的

海底水圧計で水圧変位として早期検知される津波高から、沿岸に到達する津波高を予測するモデルを構築。

方法

水圧計観測値と到達する津波高との間に、べき乗則による相関を見出し、それをもとにした非線形パラメトリックモデルを提案。

データ間のべき乗則



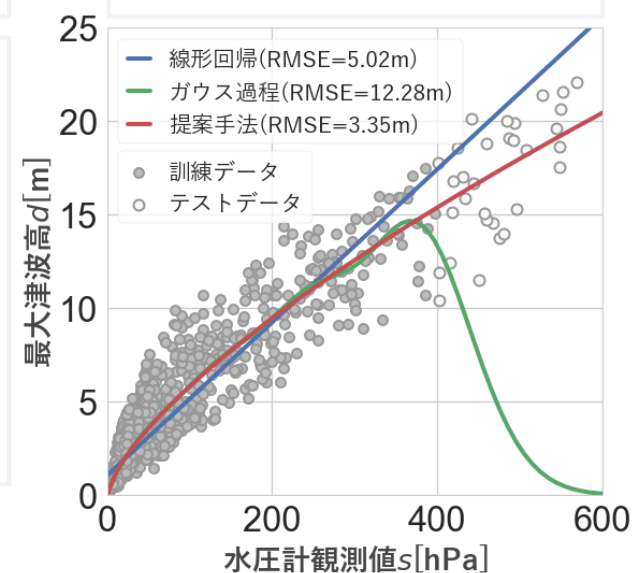
非線形パラメトリックモデル

べき乗則の式を
多変量に拡張する

$$d = as^b$$

$$\rightarrow d = \sum_i a_i s_i^{b_i}$$

外挿の予測



結果

提案手法はノンパラメトリックモデル（ガウス過程）と同等の予測精度を達成し、どの既存手法よりも外挿の精度が高くなった。

T2-10 べき乗則をもとにした津波高予測のための非線形パラメトリックモデル

吉川真史 (東大), 五十嵐康彦 (JST), 村田 伸 (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)

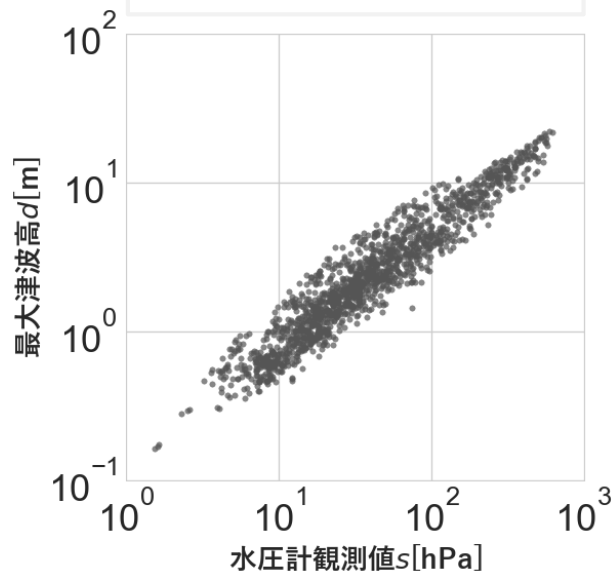
目的

海底水圧計で水圧変位として早期検知される津波高から、沿岸に到達する津波高を予測するモデルを構築。

方法

水圧計観測値と到達する津波高との間に、べき乗則による相関を見出し、それをもとにした非線形パラメトリックモデルを提案。

データ間のべき乗則



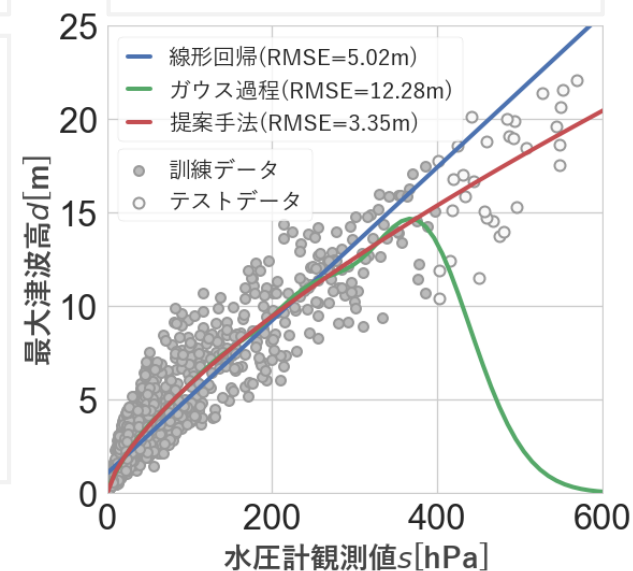
非線形パラメトリックモデル

べき乗則の式を
多変量に拡張する

$$d = as^b$$

$$\rightarrow d = \sum_i a_i s_i^{b_i}$$

外挿の予測



結果

提案手法はノンパラメトリックモデル（ガウス過程）と同等の予測精度を達成し、どの既存手法よりも外挿の精度が高くなった。

[T2-11] 多次元系列における変化点検出のための Selective Inference

梅津佑太 (名工大) ・ 竹内一郎 (名工大, 理研, NIMS)

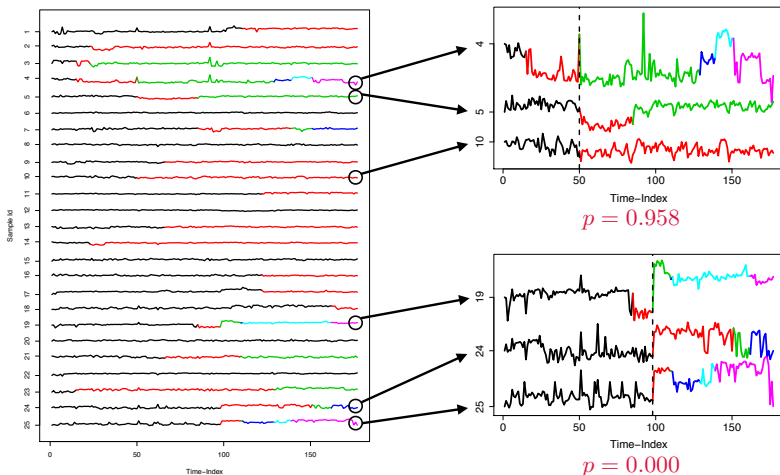
- 平均構造の変化を **Selective Inference** の枠組みで定式化
 - ✓ 検出した変化点に対して, **選択的な p -値**を付与できる
- 変化点の検出と同時に, **変化が起きた系列も検出**
- 適当な sliding window を考えることで, **多時点の変化点検出へ容易に拡張可能**

提案手法の流れ

- ① 観測された系列の疑似対数尤度比を計算
- ② 疑似対数尤度比を統合 (**aggregate**) し, 1次元の系列に変換
- ③ 変換した系列を走査 (**scan**) し, 最大化点として変化点を推定
- ④ 推定した変化点の前後で平均構造が変化したか否かを Selective Inference の枠組みで検定

実験結果

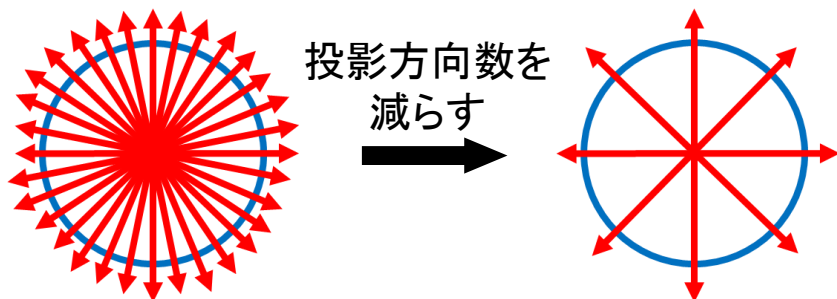
- 提案手法を Array CGH データの第 1 染色体に適用した結果
✓ 系列数 = 46, 時点数 = 177



左: 系列の一部と提案手法による変化点検出結果.
右: 検出した変化点に対応する系列と選択的な p -値.

Sparse-view X線CT

少ない投影データからCT画像の再構成を目指す



推定量

$$\hat{x} \equiv \operatorname{argmin}_x \{ \|Ax - y\|_2^2 + \lambda_1 \|\Psi x\|_1 + \lambda_2 \|x\|_{TV} \}$$

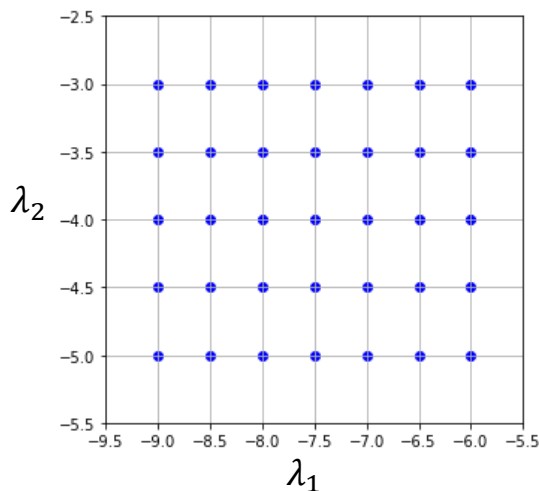
x : CT画像
 y : 投影データ
 A : 変換行列
 Ψ : wavelet変換

画像事前分布

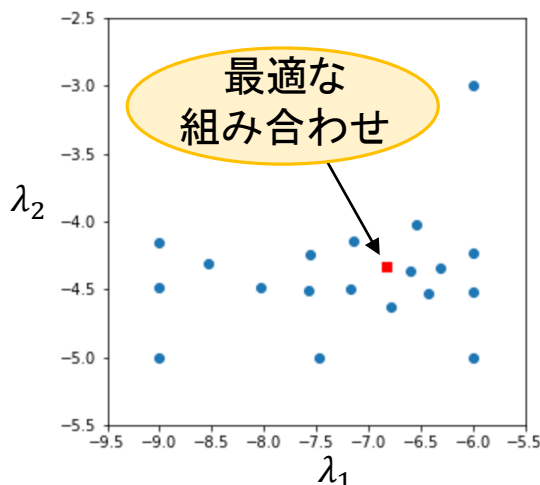
λ_1, λ_2 を上手く調整しないと
良い推定画像が得られない

提案手法

λ_1, λ_2 の組み合わせにBayes最適化を用いる



グリッドサーチ(35点)



Bayes最適化(20点)

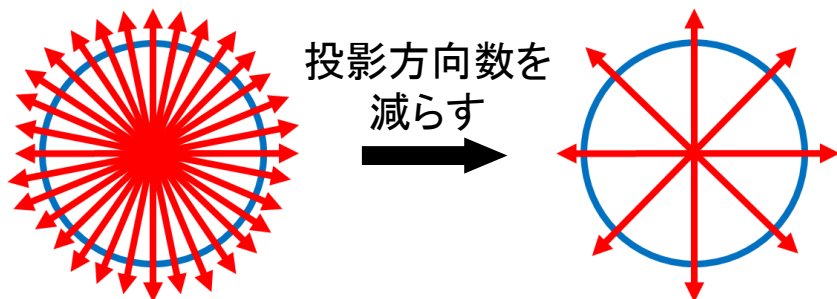
結果

- ✓ グリッドサーチに比べ、少ない試行回数で必要な部分を集中的に探索した
- ✓ 得られたハイパーパラメータは複数の肺画像に対して安定して画質を向上させた

※学生発表

Sparse-view X線CT

少ない投影データからCT画像の再構成を目指す



推定量

$$\hat{x} \equiv \operatorname{argmin}_x \{ \|Ax - y\|_2^2 + \lambda_1 \|\Psi x\|_1 + \lambda_2 \|x\|_{TV} \}$$

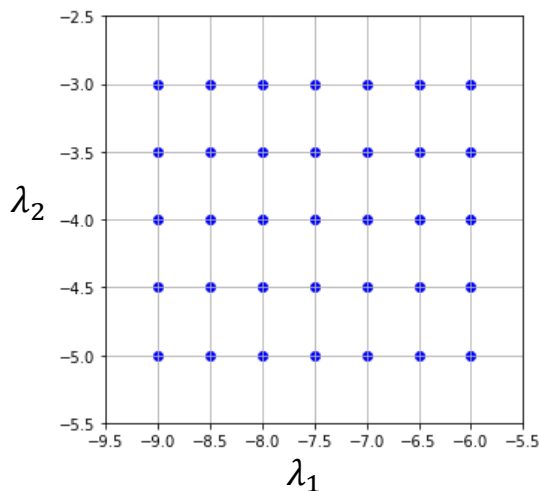
x : CT画像
 y : 投影データ
 A : 変換行列
 Ψ : wavelet変換

画像事前分布

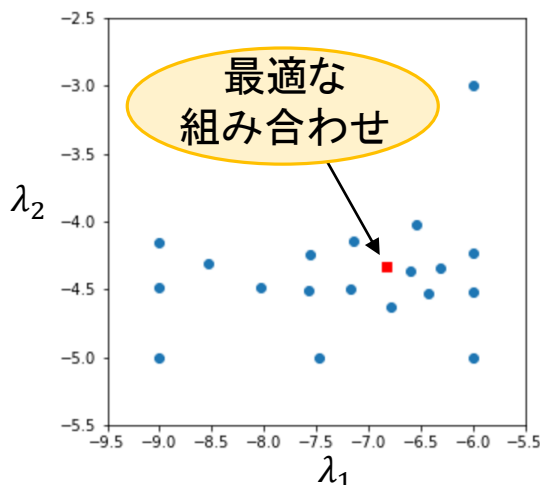
λ_1, λ_2 を上手く調整しないと
良い推定画像が得られない

提案手法

λ_1, λ_2 の組み合わせにBayes最適化を用いる



グリッドサーチ(35点)



Bayes最適化(20点)

結果

- ✓ グリッドサーチに比べ、少ない試行回数で必要な部分を集中的に探索した
- ✓ 得られたハイパーパラメータは複数の肺画像に対して安定して画質を向上させた

※学生発表

T2-14 Selective inference に基づく active learning の選択バイアス補正

稲津 佑 (理研), 竹内一郎 (名工大, 理研, NIMS)

Active learning

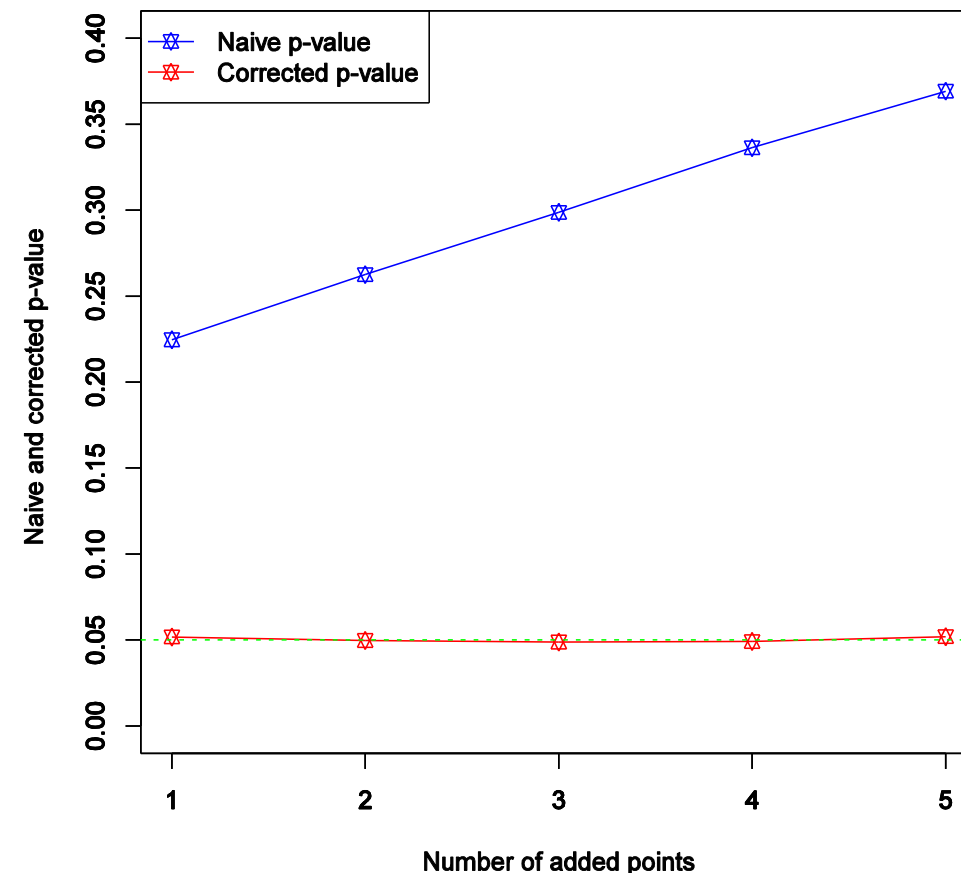
ガウス過程回帰やベイズ線形回帰等を用いた Active learning においては, 候補点を選択的に取っている

問題点

選択的に取ってきたという事実を無視して推測を行うと, 選択バイアスが生じる (e.g., 誤ったp値の評価)

解決法

- Selective inference という考え方を導入する
- 選択的に取ってきたという事象を条件付けた下での条件付き分布を用いて, 選択バイアスを補正する



ベイズ線形回帰を用いて候補点探索を行った際の, 最大値が0かどうかの検定結果 (帰無仮説が真, 有意水準0.05の場合)

T2-14 Selective inference に基づく active learning の選択バイアス補正

稲津 佑 (理研), 竹内一郎 (名工大, 理研, NIMS)

Active learning

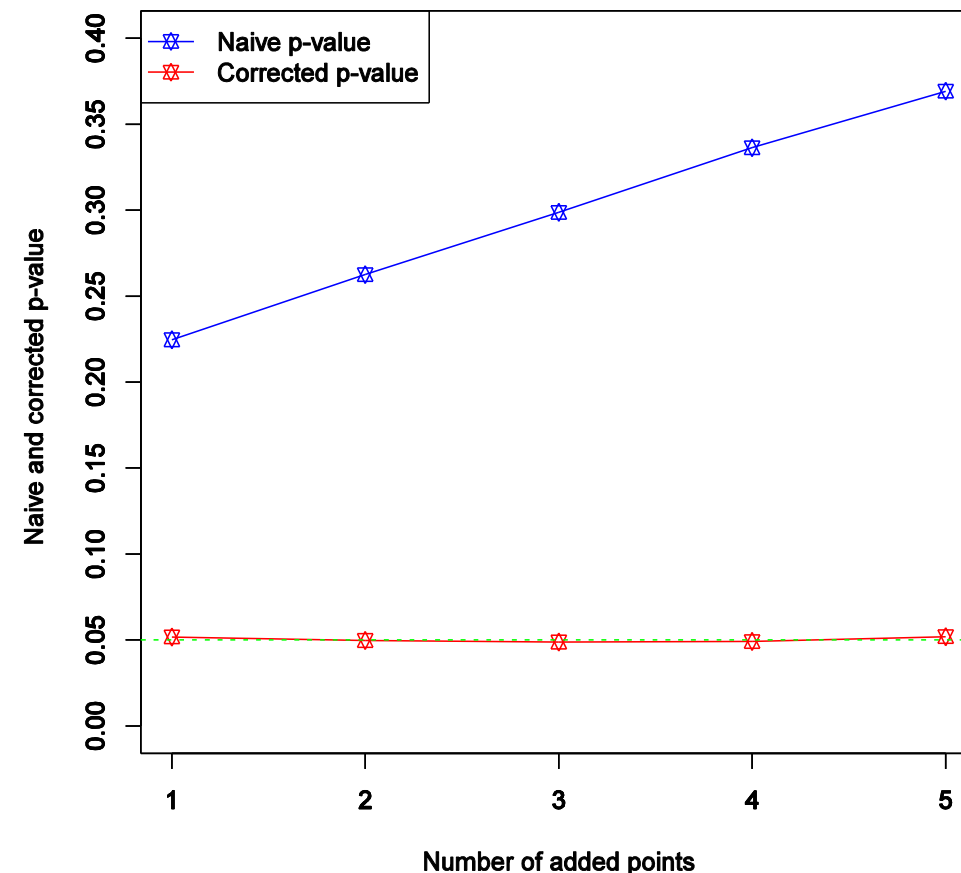
ガウス過程回帰やベイズ線形回帰等を用いた Active learning においては, 候補点を選択的に取っている

問題点

選択的に取ってきたという事実を無視して推測を行うと, 選択バイアスが生じる (e.g., 誤ったp値の評価)

解決法

- Selective inference という考え方を導入する
- 選択的に取ってきたという事象を条件付けた下での条件付き分布を用いて, 選択バイアスを補正する



ベイズ線形回帰を用いて候補点探索を行った際の, 最大値が0かどうかの検定結果 (帰無仮説が真, 有意水準0.05の場合)

T2-15 2つの疎な行列からなる行列分解問題の変分ベイズ法による解析法の検討

玉井 智貴 竹田 晃人（茨城大学）

学生優秀プレゼンテーション賞対象

目的

- 観測行列：2つの疎な非負値行列の積
- ⇒ 行列分解問題の解析解を求める
- 解析解を数値実験で評価

モデル

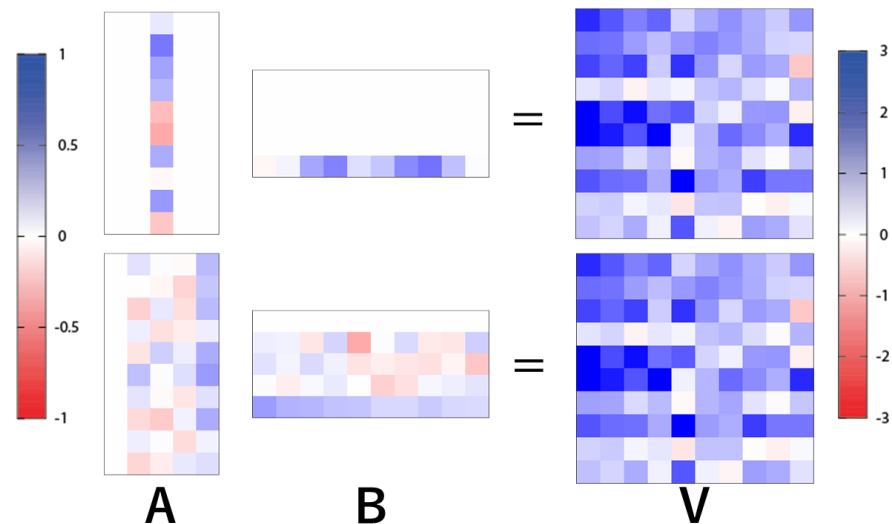
$$\text{観測行列 } V = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} A \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} B + \text{ノイズ行列 } E$$

- 事前分布：指数分布
指数分布のパラメータは λ （疎性を表す）
- モデル尤度：ノイズ行列がガウス分布

⇒変分ベイズ法で解析解を求める

数値実験結果

上が $\lambda = 1$,下が $\lambda = 100$ の結果



- 分解行列の疎性に λ が影響を与えている
- 分解行列は非負値行列にならなかった

T2-15 2つの疎な行列からなる行列分解問題の変分ベイズ法による解析法の検討

玉井 智貴 竹田 晃人 (茨城大学)

学生優秀プレゼンテーション賞対象

目的

- 観測行列：2つの疎な非負値行列の積
- ⇒ 行列分解問題の解析解を求める
- 解析解を数値実験で評価

モデル

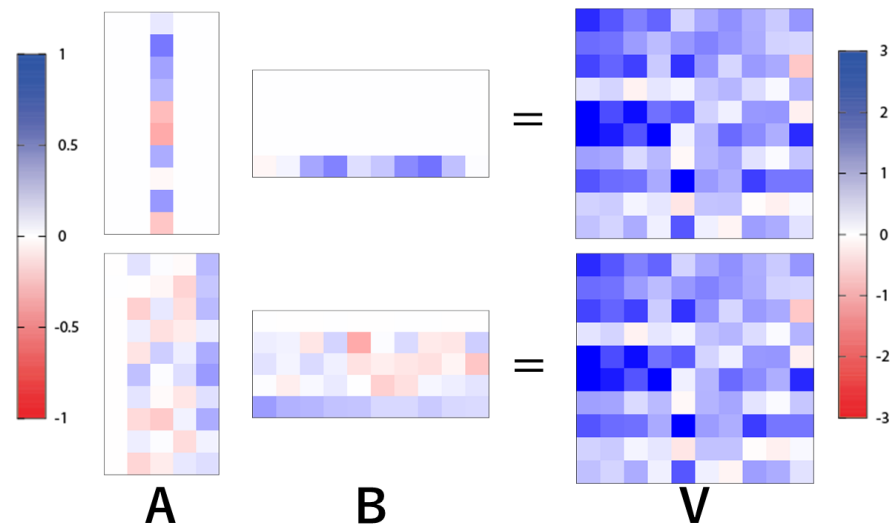
$$\text{観測行列 } V = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} A \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} B + \text{ノイズ行列 } E$$

- 事前分布：指数分布
指数分布のパラメータは λ (疎性を表す)
- モデル尤度：ノイズ行列がガウス分布

⇒変分ベイズ法で解析解を求める

数値実験結果

上が $\lambda = 1$,下が $\lambda = 100$ の結果

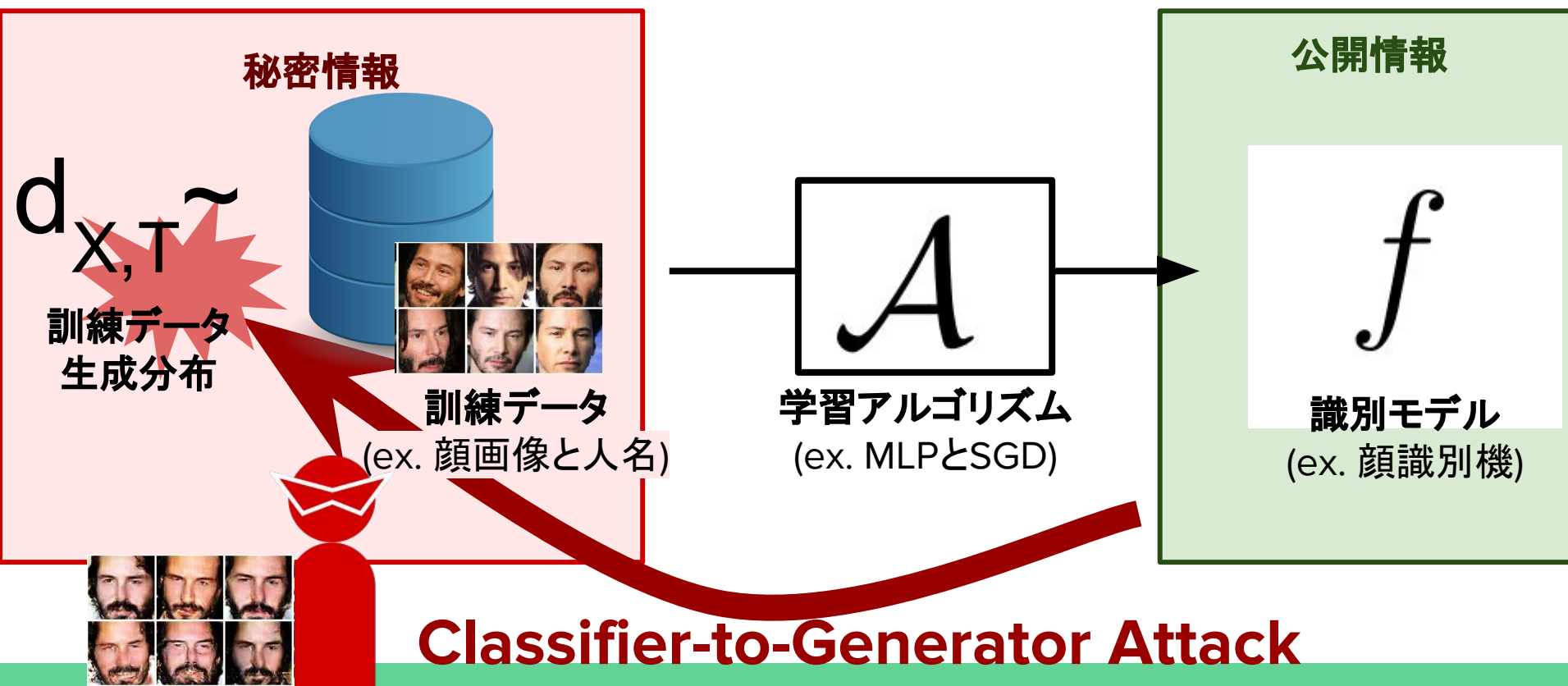


- 分解行列の疎性に λ が影響を与えている
- 分解行列は非負値行列にならなかった

Generative Adversarial Networksを用いた 確率的識別モデルから訓練データ生成分布の推定

草野 光亮 (筑波大) / 佐久間 淳 (筑波大/理研AIP)

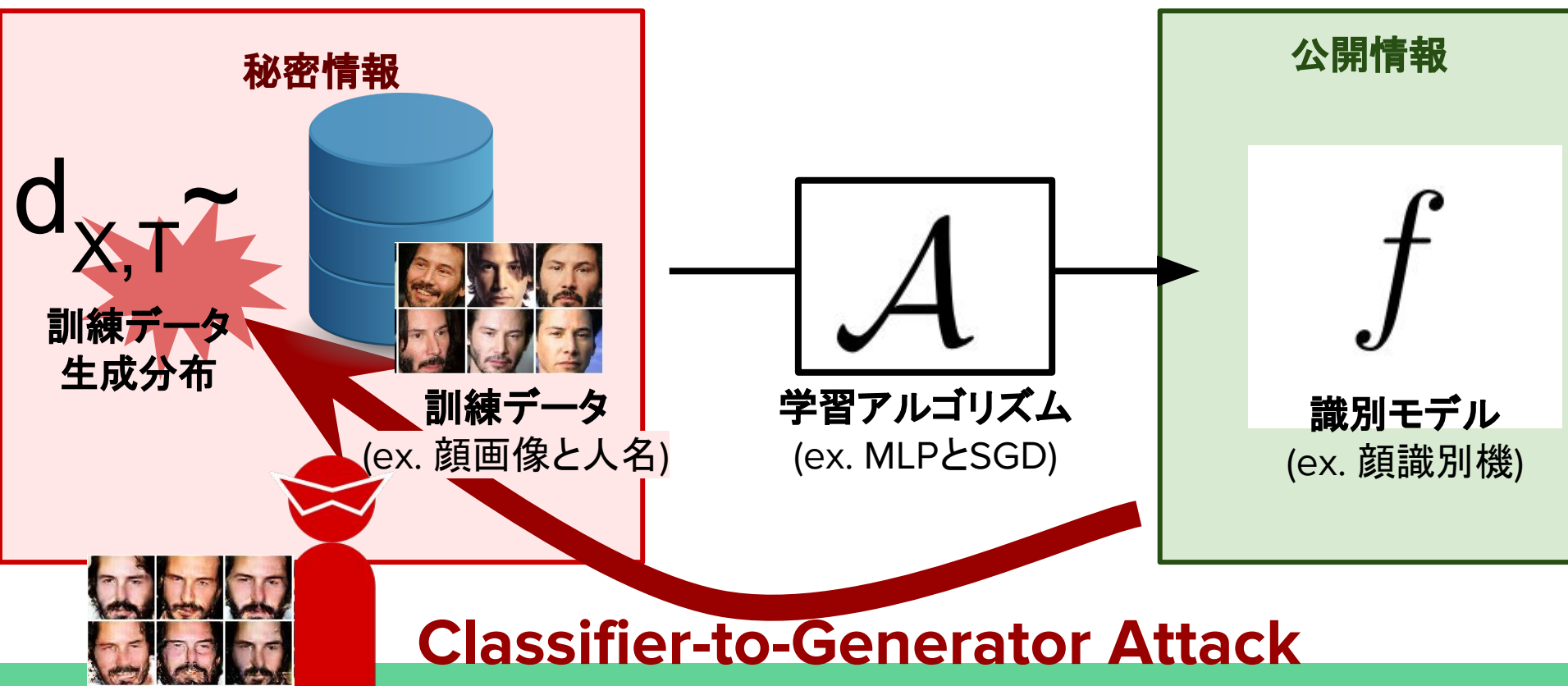
- **訓練データとその生成分布**が秘密情報である場合、学習により得られた**識別モデル**から、秘密情報である**訓練データ生成分布**を推定できるか？
- Generative Adversarial Networksを利用した攻撃アルゴリズム**PrelmgaGAN**を提案
- 実験により、識別モデルから訓練データ生成分布を**推定可能**であることを示した。



Generative Adversarial Networksを用いた 確率的識別モデルから訓練データ生成分布の推定

草野 光亮 (筑波大) / 佐久間 淳 (筑波大/理研AIP)

- **訓練データとその生成分布**が秘密情報である場合、学習により得られた**識別モデル**から、秘密情報である**訓練データ生成分布**を推定できるか？
- Generative Adversarial Networksを利用した攻撃アルゴリズム**PrelmgaGAN**を提案
- 実験により、識別モデルから訓練データ生成分布を**推定可能**であることを示した。



T2-17: Tree-reweighted近似による Ising逆問題の解

佐野崇 産総研AI研究センター

背景

- Ising模型; 2値変数マルコフ確率場
 - シンプルだが推論(順問題)も学習(逆問題)も困難
- Bethe近似による逆問題の解析解が先行研究で求められている

$$P(\mathbf{s}) = \frac{1}{Z} \exp(-E(\mathbf{s})) \quad E(\mathbf{s}) = - \sum_{\langle ij \rangle \in E} J_{ij} s_i s_j - \sum_{i \in V} h_i s_i$$

寄与した点

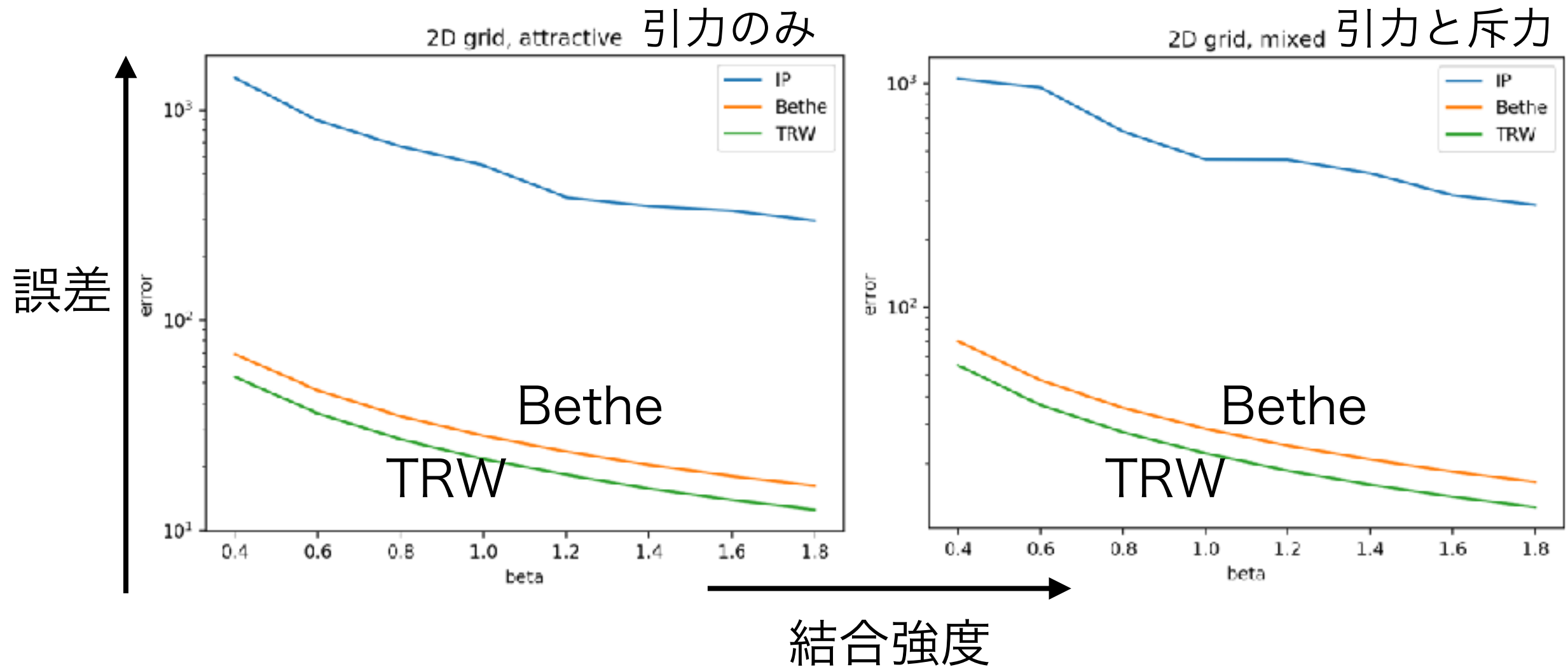
- Tree-ReWeighted(TRW)近似による逆問題の解を求めた
 - TRW近似自由エネルギーを最小化するパラメータを、データの相関行列の関数として解析的に与えた
- 反復計算をせずにパラメータが推定できる
- 真の目的関数の下限を最大化(Bethe近似にはない利点)

$$J_{ij}^{TRW} = -\rho_{ij} a \tanh \left[\frac{1}{2(\tilde{C}^{-1})_{ij}} \sqrt{\tilde{D}_{ij}} \dots \right]$$

C: 相関行列

数値評価

- パラメータを固定したIsing模型からGibbsサンプリングで相関行列を計算
- TRW近似、Bethe近似の逆問題の解に相関行列を代入して、真のパラメータとの誤差を評価

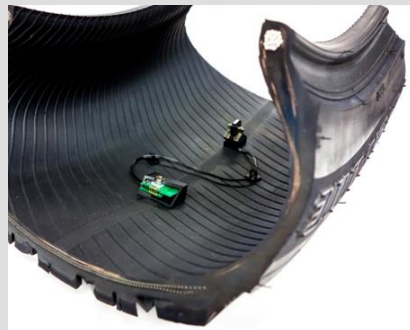


結果

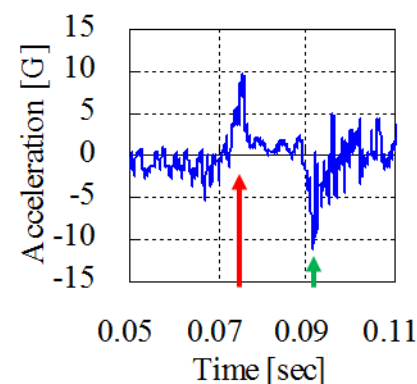
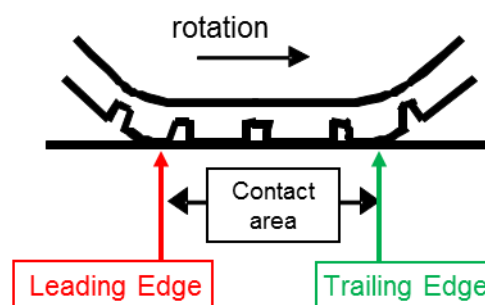
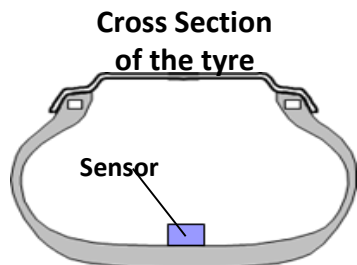
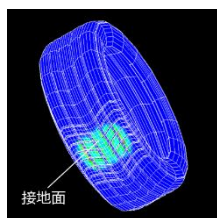
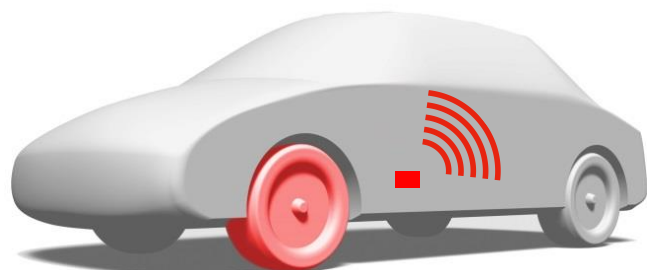
- TRW近似はBethe近似より精度良くパラメータを推定できた
- TRW近似が真の目的関数の下限を最大化している効果

T2-18: EMDを用いたタイヤセンシングのための特徴抽出法

○石井 啓太（ブリヂストン），後藤 嵩人（ブリヂストン），松井 知子（統計数理研），
Gareth W. Peters（HWU），Nourddine Azzaoui（UCA）



- ◆ タイヤ内側に加速度センサーを貼付し、**振動波形から路面状態を分類**
- ◆ 特徴抽出に**Empirical Mode Decomposition**を利用
- ◆ EMDが**瞬時的な波形の周波数解析が可能**であるため、タイヤ振動の解析に有用
- ◆ 従来法と比較して、**判定精度を維持したまま計算コストを大幅に低減**



背景

- ・教師なし異常検知において与えられる正常時データの詳細なラベル(運転モード等)は不明
- ・近傍法では正常時データの不均衡さが性能に悪影響を与えることがある

目的

正常時データの不均衡さの解消

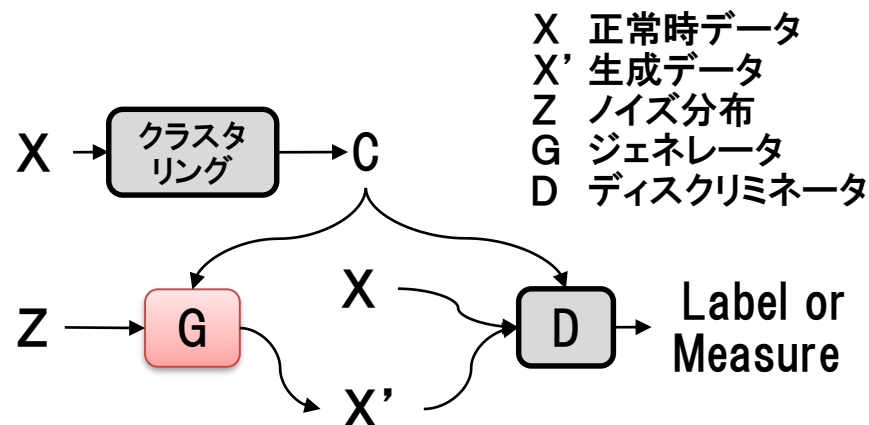
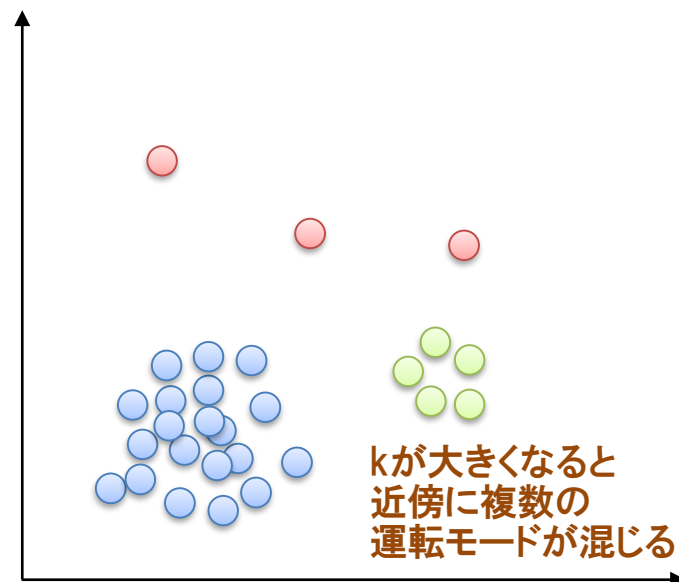
手法

クラスタリング **×** Conditional GAN

によるオーバー&アンダーサンプリング

➡ 性能(上位N件精度)の向上を確認

- 正常時データ(運転時)
- 正常時データ(立上げ/立下げ時)
- 異常時データ



背景

- ・教師なし異常検知において与えられる正常時データの詳細なラベル(運転モード等)は不明
- ・近傍法では正常時データの不均衡さが性能に悪影響を与えることがある

目的

正常時データの不均衡さの解消

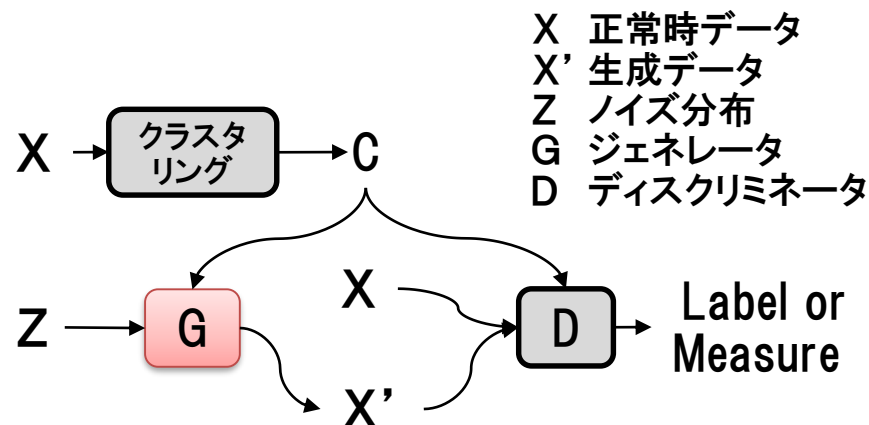
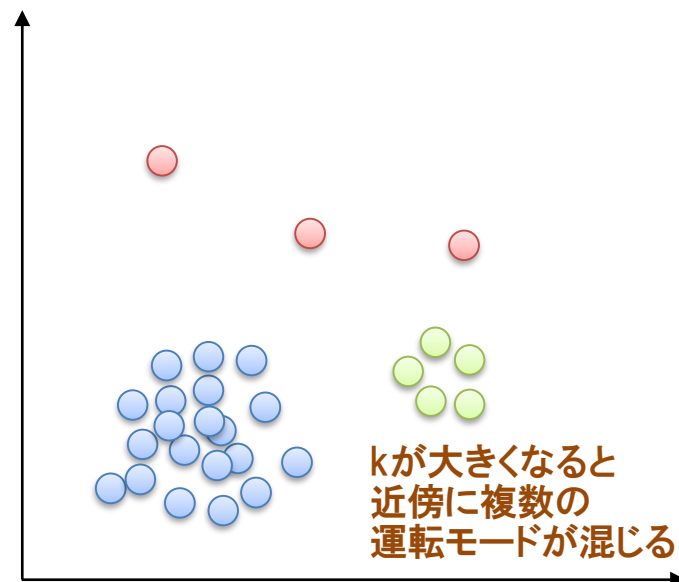
手法

クラスタリング **×** Conditional GAN

によるオーバー&アンダーサンプリング

➡ 性能(上位N件精度)の向上を確認

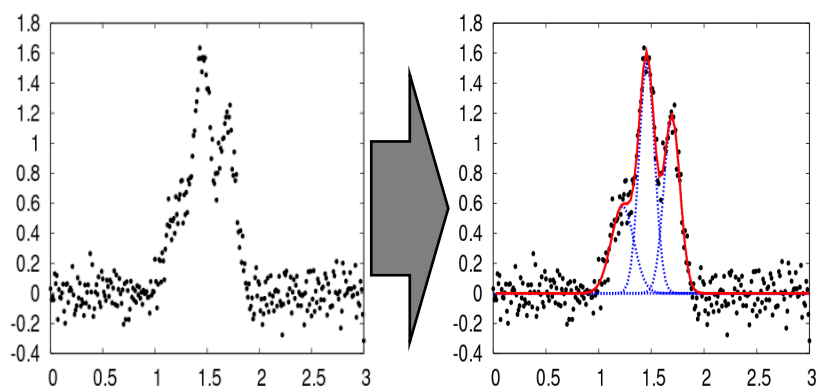
- 正常時データ(運転時)
- 正常時データ(立上げ/立下げ時)
- 異常時データ



スペクトル分解における λ -スキャン法の提案

本武陽一(東大), 五十嵐康彦(JSTさがけ),
竹中 光(東大), 永田賢二(産総研), 岡田真人(東大)

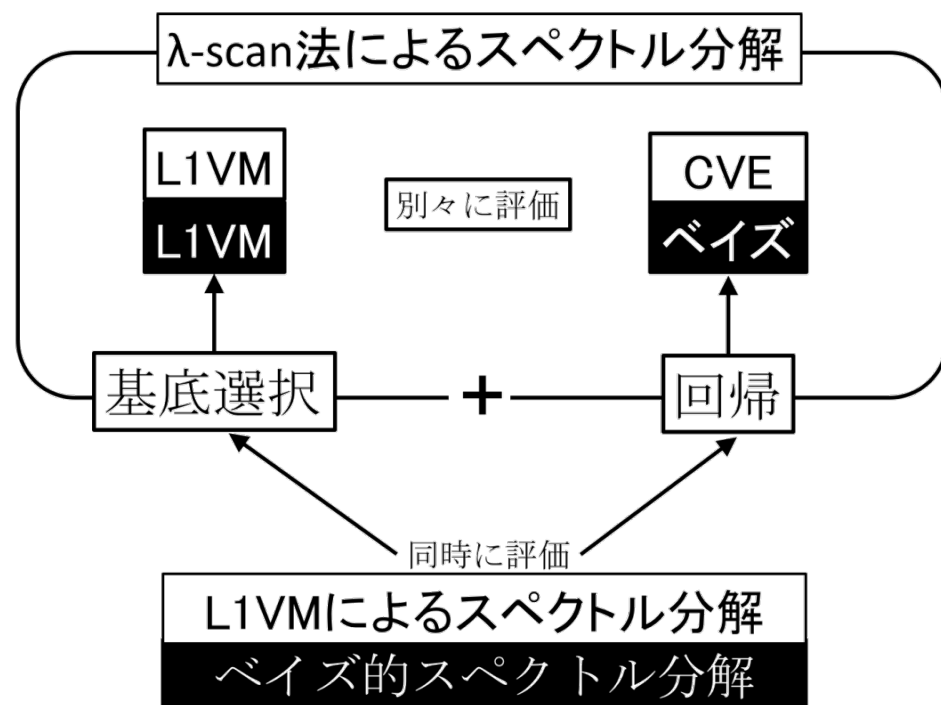
◆スペクトル分解



▷ スペクトルデータから
ピーク数・位置・強度を抽出

◆提案手法: λ -scan法

- ▷ スペクトル分解 = 基底探索 + 回帰
- ▷ ベイズ的自由エネルギーが解析的に求まる

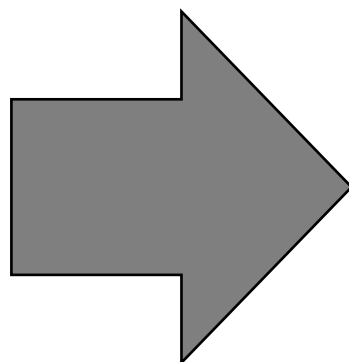
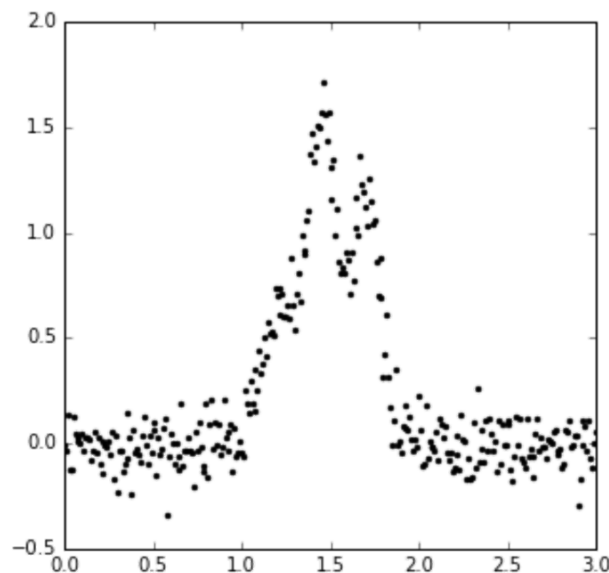


目的: 高速・高性能なスペクトル分解を実現

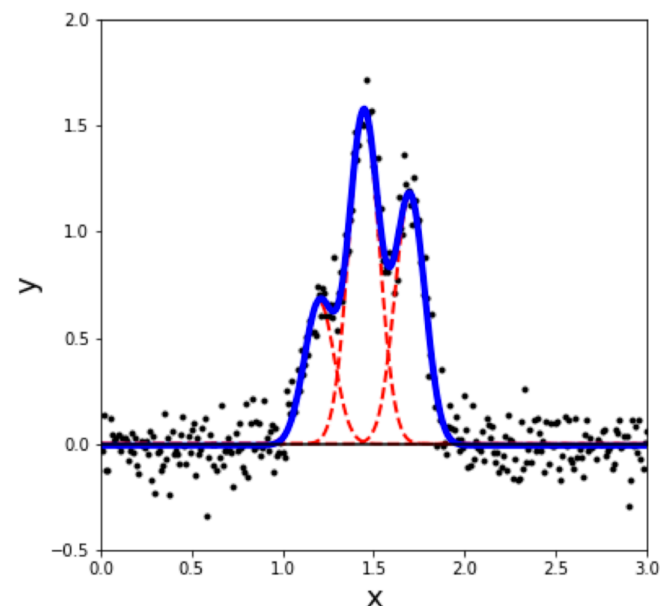
スペクトル分解における λ -スキャン法の提案

本武陽一(東大), 五十嵐康彦(JSTさがけ),
竹中 光(東大), 永田賢二(産総研), 岡田真人(東大)

◆スペクトル分解結果



λ -scan法による
スペクトル分解



アルゴリズム	計算時間:3peak	計算時間:8peak
λ -scan 法によるスペクトル分解	12.03 [min.]	10.8 [min.]
[既存手法] ベイズ的スペクトル分解	59.84 [min.]	89.42 [min.]

結論: 高速・高性能なスペクトル分解を実現

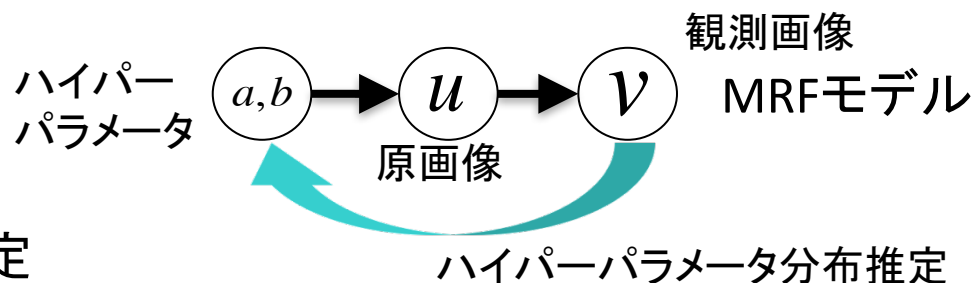
T2-21 ガウス過程とベイズ最適化を用いた近似的ハイパーパラメータ分布推定

片上 舜, 坂本 浩隆, 岡田 真人(東大)

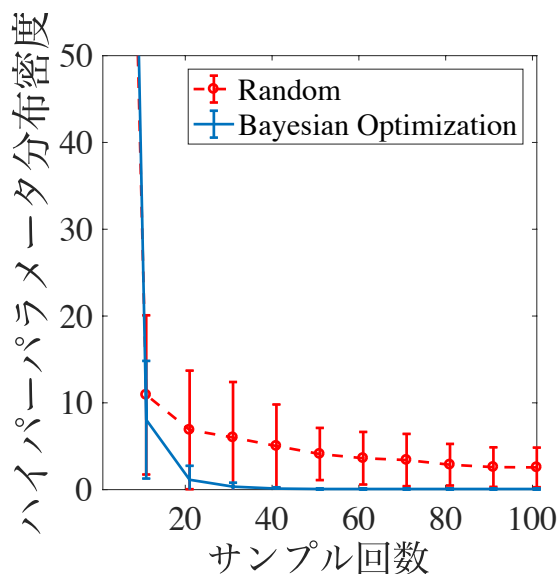
学生優秀プレゼンテーション賞対象

目的

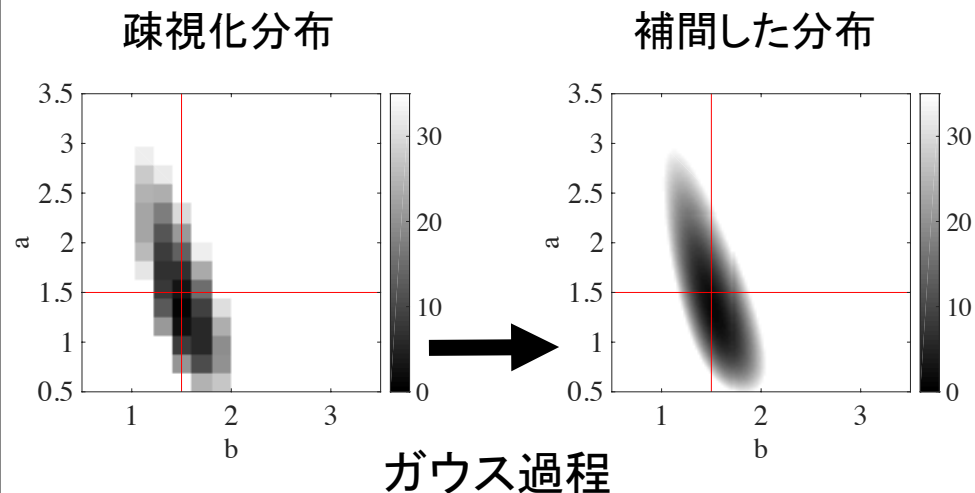
ガウス過程とベイズ最適化を用いた
画像のハイパーパラメータ高速分布推定



ハイパーパラメータ点推定



ハイパーパラメータ分布推定



結果

提案手法により, 少数サンプリングからハイパーパラメータ分布推定および,
ハイパーパラメータの点推定が可能であることが確認された.

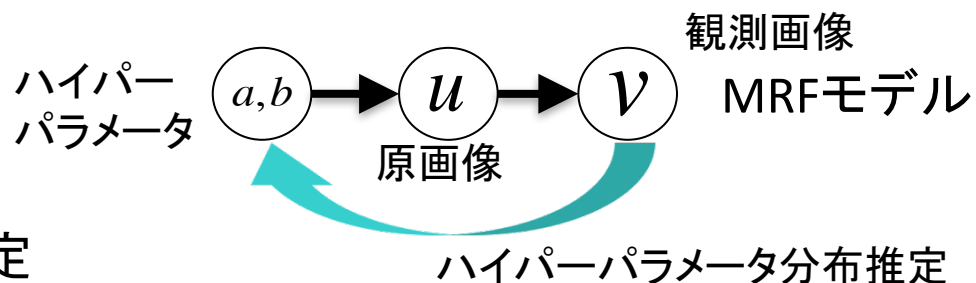
T2-21 ガウス過程とベイズ最適化を用いた近似的ハイパーパラメータ分布推定

片上 舜, 坂本 浩隆, 岡田 真人(東大)

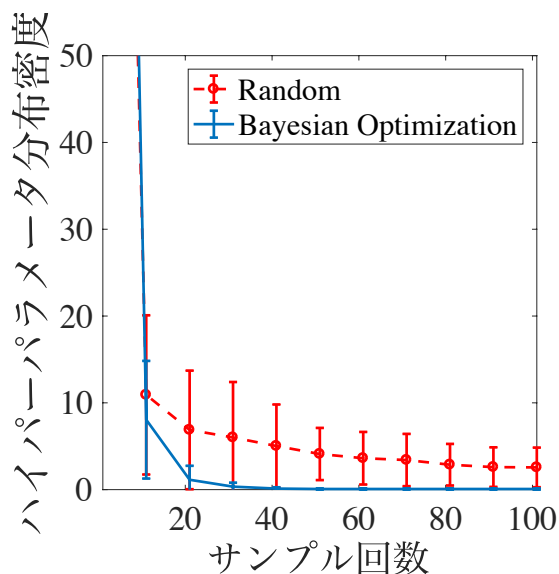
学生優秀プレゼンテーション賞対象

目的

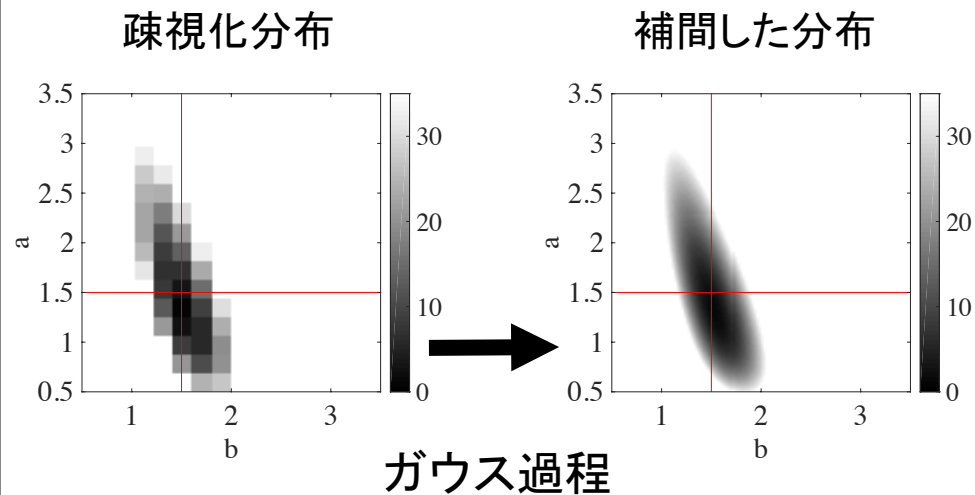
ガウス過程とベイズ最適化を用いた
画像のハイパーパラメータ高速分布推定



ハイパーパラメータ点推定



ハイパーパラメータ分布推定



結果

提案手法により, 少数サンプリングからハイパーパラメータ分布推定および,
ハイパーパラメータの点推定が可能であることが確認された.

鹿野 英明^{1,2} 本多 淳也^{1,2} 坂巻 顕太郎¹ 松浦 健太郎³ 中村 篤祥⁴ 杉山 将^{2,1}
¹東京大学 ²理化学研究所 ³Johnson & Johnson ⁴北海道大学

オンライン広告や医薬品開発では

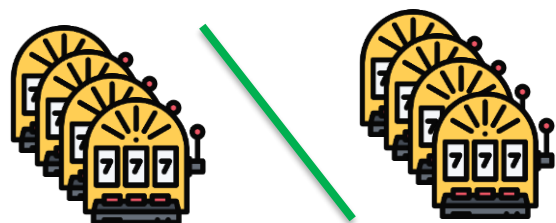
満足*のいく*効果が得られるものをできるだけ早く見つけることが重要

優良腕：期待値がしきい値以上の台

従来法：しきい値バンディット

しきい値未満

しきい値以上



全ての台を優良腕と
それ以外の台に分類

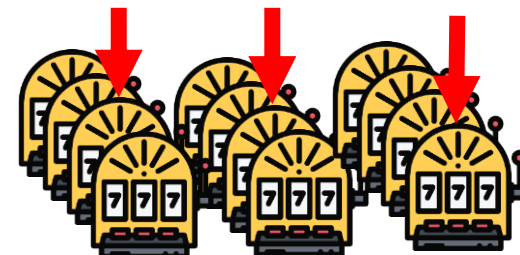
全ての台の
探索回数がそれぞれ

$$O\left(\log \frac{1}{\delta}\right)$$

実用の場面では
全ての台を分類する
必要はなく優良腕を
必要な精度(誤り確率 δ)で
必要な数だけ特定できれば
十分なのでは？

新たな枠組み：優良腕識別

Good Good Good



優良腕 λ 台を見つけるときの
探索回数を最小化する

上位 λ 個以外の台の
探索回数はそれぞれ

$$O\left(\log \log \frac{1}{\delta}\right)$$

既存法は悪い台も多く探索 ☹️

提案法は当たりやすい台を
優先的に探索 😊

鹿野 英明^{1,2} 本多 淳也^{1,2} 坂巻 顕太郎¹ 松浦 健太郎³ 中村 篤祥⁴ 杉山 将^{2,1}
¹東京大学 ²理化学研究所 ³Johnson & Johnson ⁴北海道大学

オンライン広告や医薬品開発では

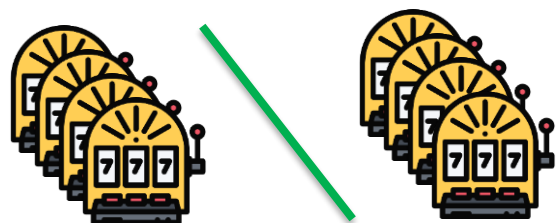
満足*のいく*効果が得られるものをできるだけ早く見つけることが重要

優良腕：期待値がしきい値以上の台

従来法：しきい値バンディット

しきい値未満

しきい値以上



全ての台を優良腕と
それ以外の台に分類

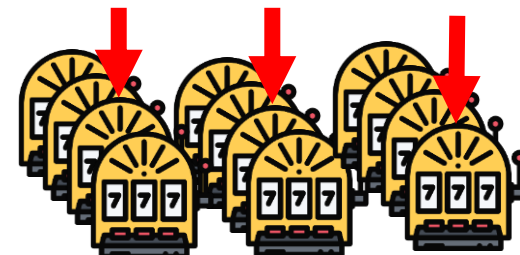
全ての台の
探索回数がそれぞれ

$$O\left(\log \frac{1}{\delta}\right)$$

実用の場面では
全ての台を分類する
必要はなく優良腕を
必要な精度(誤り確率 δ)で
必要な数だけ特定できれば
十分なのでは？

新たな枠組み：優良腕識別

Good Good Good



優良腕 λ 台を見つけるときの
探索回数を最小化する

上位 λ 個以外の台の
探索回数はそれぞれ

$$O\left(\log \log \frac{1}{\delta}\right)$$

既存法は悪い台も多く探索 ☹️

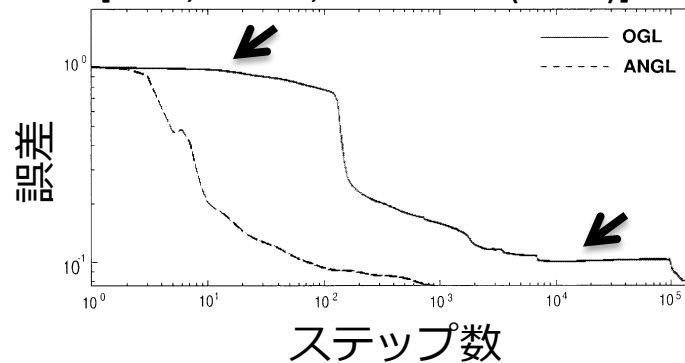
提案法は当たりやすい台を
優先的に探索 😊

複数の出力素子をもつ三層パーセプトロンの 学習ダイナミクスの統計力学的解析

～ プラトー現象の再検討 ～

吉田雄紀^{1*}, 唐木田亮², 岡田真人^{1,2,3}, 甘利俊一³ (¹東大, ²産総研, ³理研)

[Park, Amari, Fukumizu (1999)]

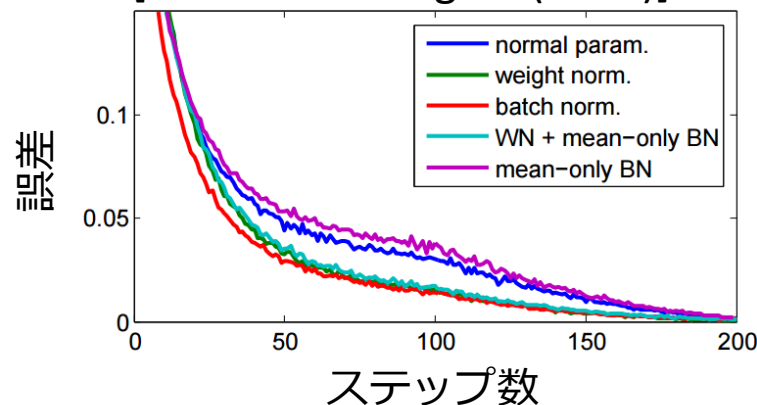


プラトー：学習の停滞

90年代後半

「プラトーに阻まれ学習困難」

[Salimans & Kingma (2016)]



近年の深層学習では
プラトーを見ることは稀

→ **プラトー理論の再検討**

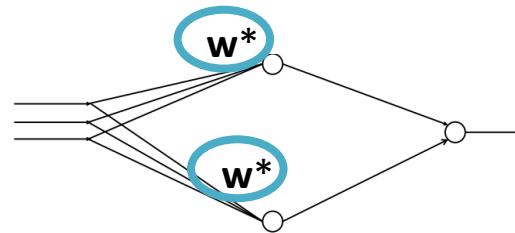
複数の出力素子をもつ三層パーセプトロンの 学習ダイナミクスの統計力学的解析

～ プラトー現象の再検討 ～

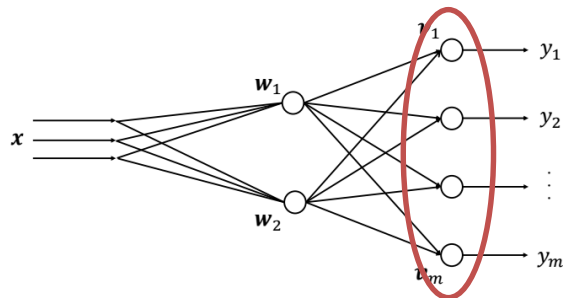
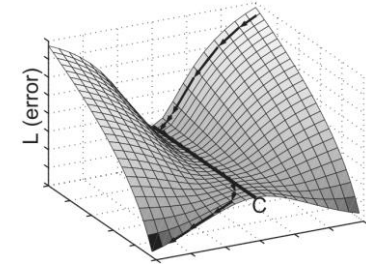
吉田雄紀^{1*}, 唐木田亮², 岡田真人^{1,2,3}, 甘利俊一³ (1東大, 2産総研, 3理研)

プラトー理論

特異領域がもたらす
ミルナーアトラクタにより
学習が停滞



[Fukumizu & Amari (2000)]



これまでのプラトー理論研究は
出力が1次元

[Saad & Solla (1995), Riegler & Biehl (1995),
Fukumizu & Amari (2000), Wei et al.(2008),
Cousseau et al.(2008)]

→ 複数次元出力の場合
にプラトー理論を再検討

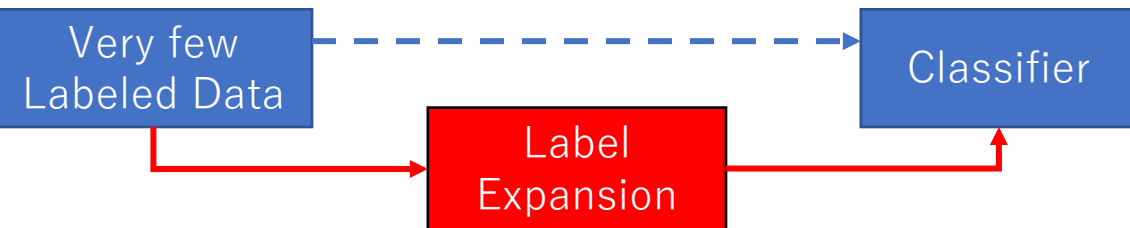
ラベル拡張を考慮した劣モジュラ性に基づく能動学習

木村正成¹⁾, 若林啓²⁾

¹筑波大学 情報学群 知識情報・図書館学類 ²筑波大学 図書館情報メディア系

概要

- 新規プロジェクトなどで教師データが極端に少ないケースで、少ないアノテーションで良い分類器を学習する研究.



貢献

- ラベル付きデータが希少な場合に有効なラベル拡張の手法と、それを能動学習と組み合わせた新しい手法を提案.
- ラベル拡張を考慮した能動学習の手法として、グラフ構造に基づいた手法を提案.
- 実際のデータセットに対して実験を行い提案手法の有用性を示す.

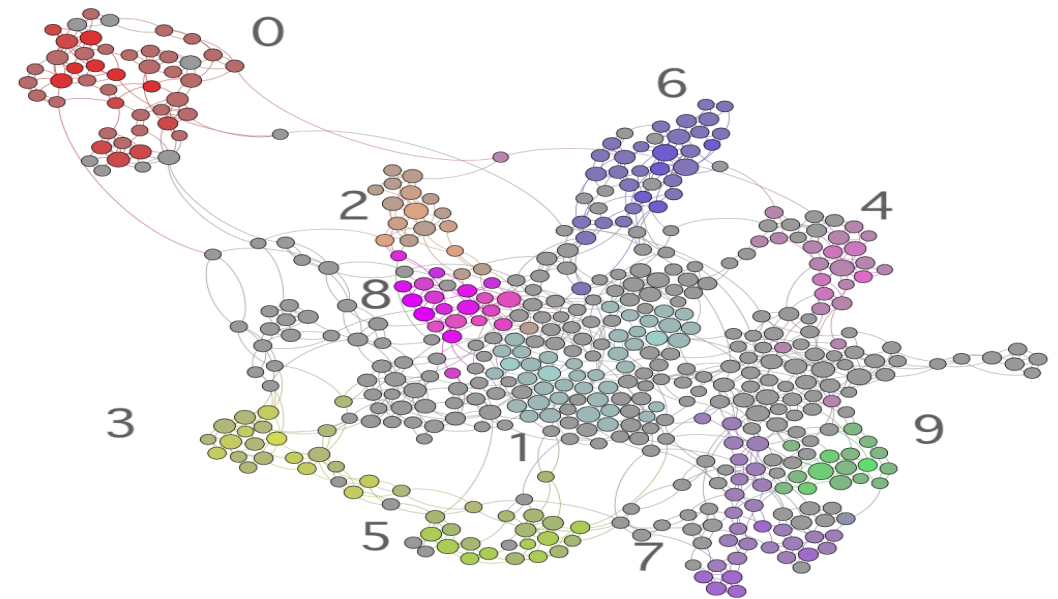
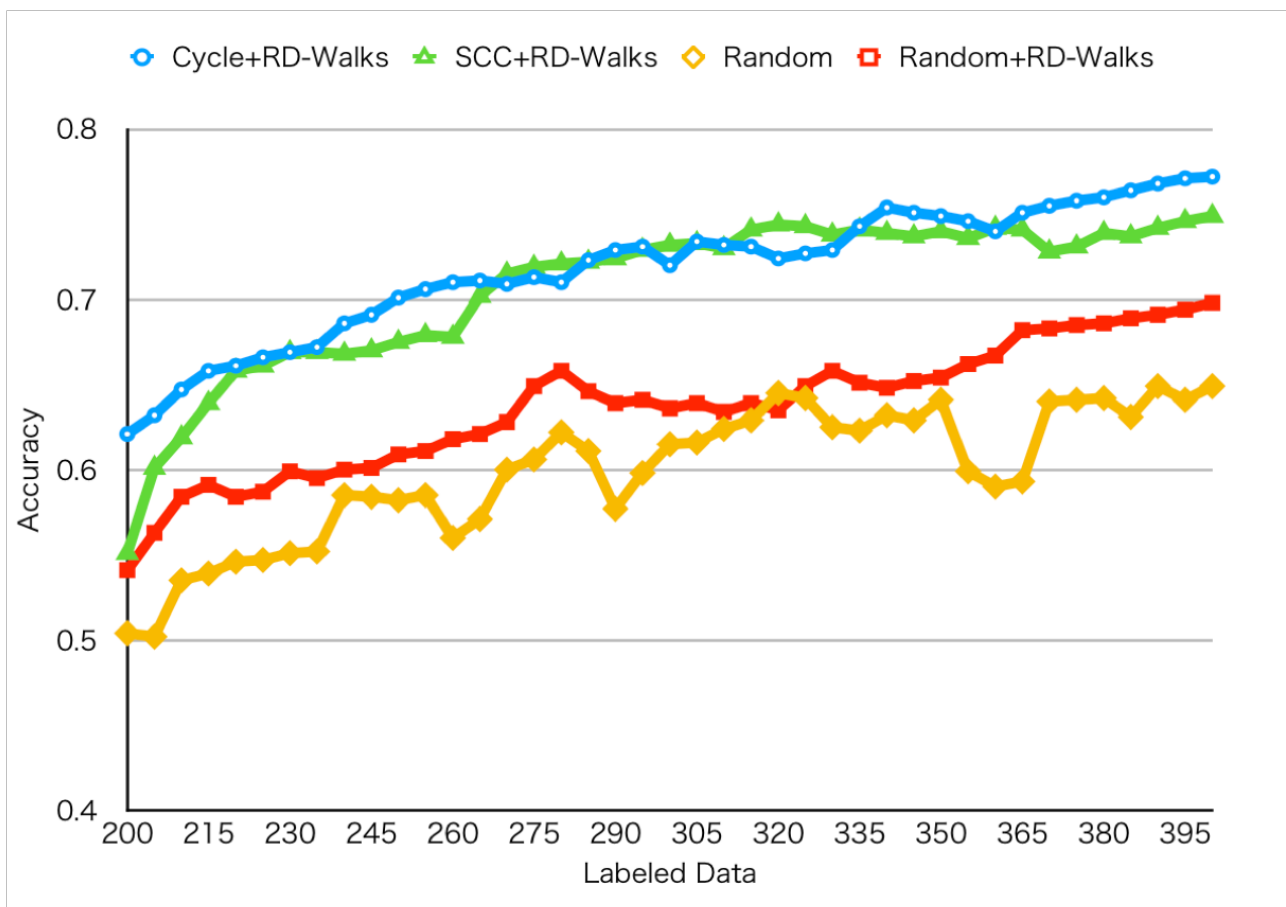


Fig1. MNISTに対して1-shotでラベル拡張を行なった図. このようなラベルの拡張範囲を最大化する最大被覆問題として捉え、能動学習を行う.

実際のデータセットに対する実験結果



◆, ▲が提案手法を用いた結果

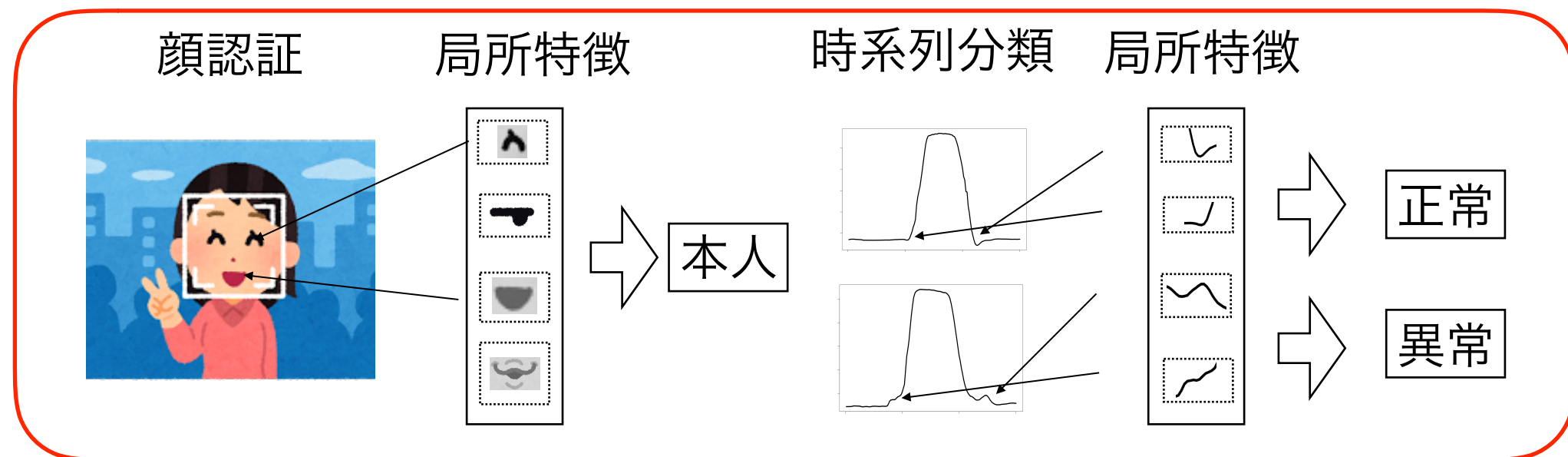
- 縦軸がAccuracy
- 横軸がラベル付きデータ数

ラベル拡張及び能動学習の手法によって少ないデータ数からでも良い分類器を構築できていることがわかる

Learning theory and algorithms for shapelets and other local features

末廣 大貴, 畑埜 晃平 (九大 / AIP), 瀧本 英二 (九大),
山本 修司, 坂内 健一 (慶大 / AIP), 武田 朗子 (統数研 / AIP)

局所特徴を使った分類手法は非常にポピュラー



Q. 「良い局所特徴」をどうやって学習するか？

➡ 実は統一的な枠組みがない (ドメイン依存)
理論的な汎化性能保証がない

主結果

■ 局所特徴学習に対する

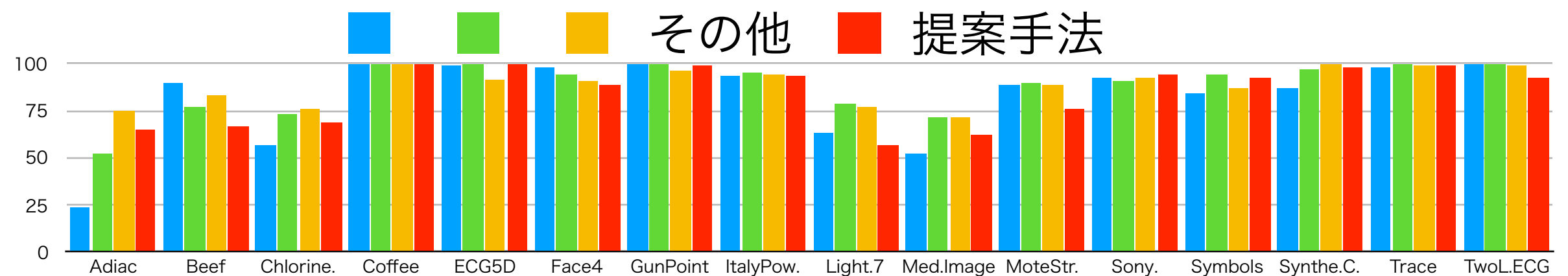
◆ 統一的な枠組み

◆ アルゴリズム

◆ 汎化性能保証

全てを与えた

■ 最新手法に匹敵する実験精度



T2-26 対数線形モデルの学習における L_1 正則化後の最尤推定の効果

通常の L_1 正則化は
構造学習) パラメータ削減によるモ
デル多様体の決定
パラメータ学習) 削減後のパラメー
タの調節
を同時に行います.

では「構造学習だけを L_1 正則化で行いパラメータ学習は最尤推定で行う」というオプションはどうなるでしょうか？

元々の L_1 正則化のみの場合に比べて性能は良くなるでしょうか？

比較結果はポスターで.

T2-27: 一般化ラベルノイズモデルにおける分類問題について

須子統太(早大), 堀井俊佑(早大)

分類問題において
様々なノイズ混入モデルを統一的
に表現するモデルを提案

⇒EM法とVB法による分類アルゴリズムを構成し数値実験

本来とは異なるラベルが得られる
(⇒**正例と負例が反転するノイズ**)

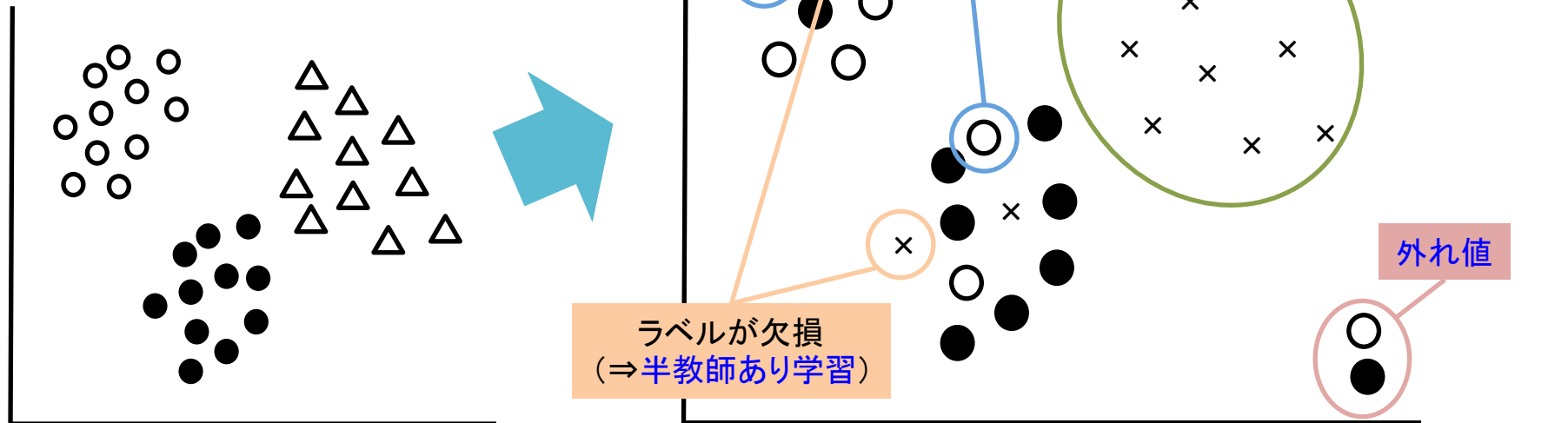
あるラベルは全て欠損
(⇒**正例とラベルなしデータからの学習**)

ラベルが欠損
(⇒**半教師あり学習**)

外れ値

理想的なデータ

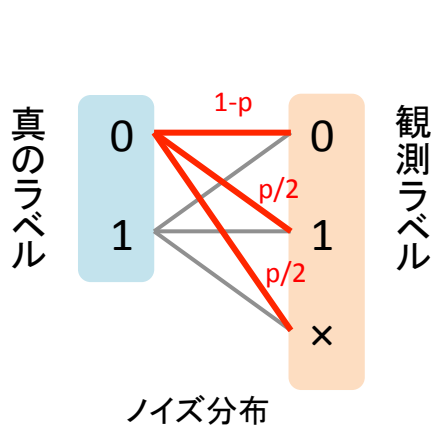
提案ノイズモデルで表現可能なデータの例



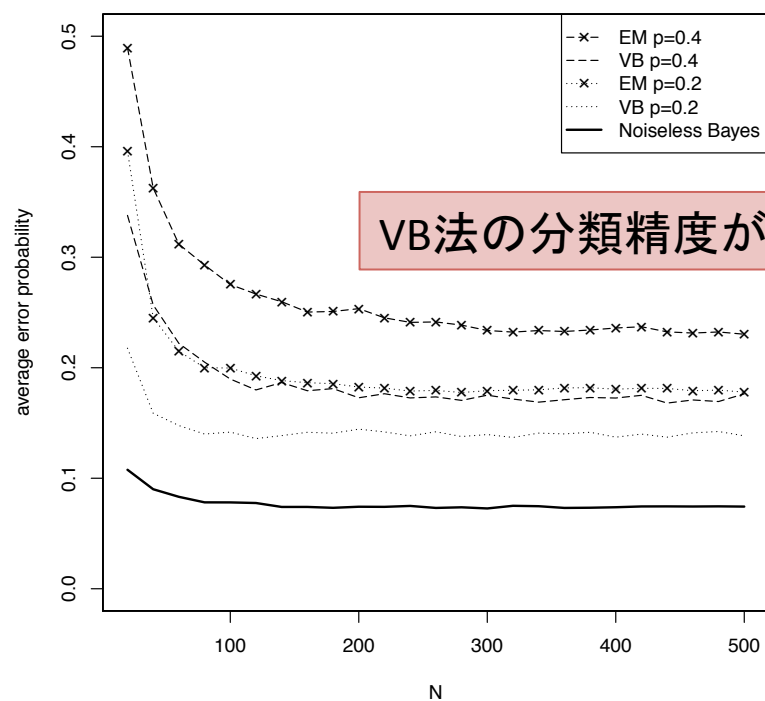
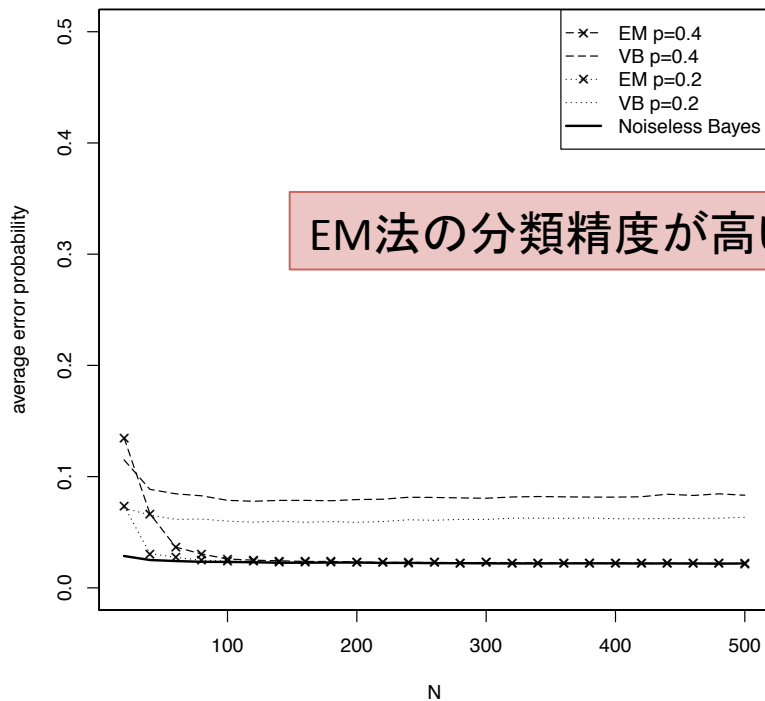
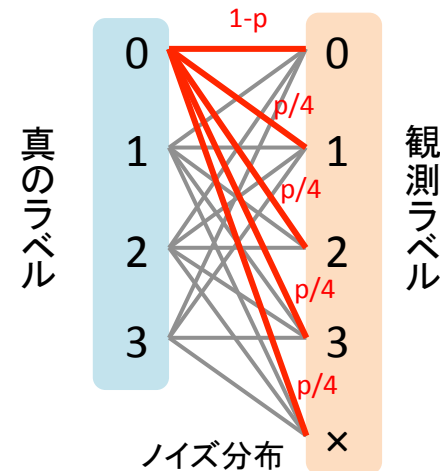
T2-27: 一般化ラベルノイズモデルにおける分類問題について

須子統太(早大), 堀井俊佑(早大)

◆ 真のラベル数=2の場合



◆ 真のラベル数=4の場合



Graph Product Multilayer Network型 Gauss Markov確率場の構造学習

高品佑也（早大）・井上真郷（早大）

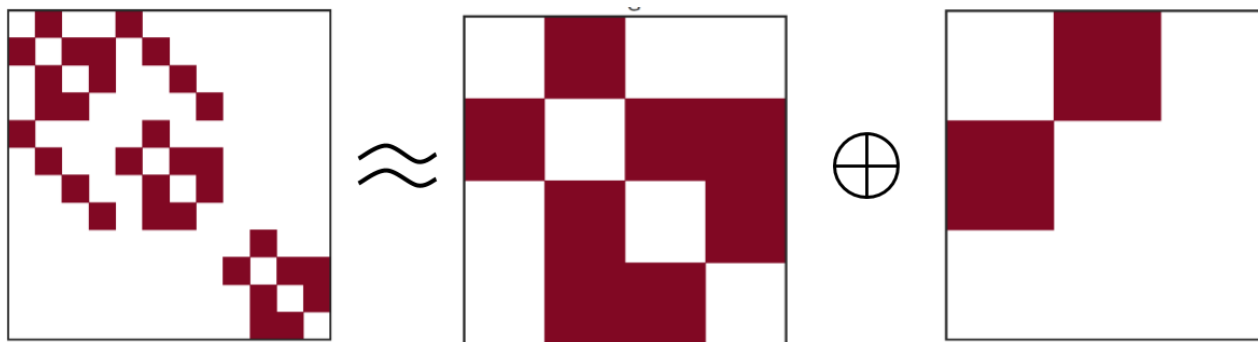
Gauss Markov確率場の構造推定

- ・ グラフ構造に対応する精度行列 Θ を MAP 推定。

$$\underset{\Theta}{\operatorname{argmax}} \{ \ln \det \Theta - \operatorname{tr}(\mathbf{S}\Theta) - \rho \|\Theta\|_1 \}$$

本研究: 全体のグラフを階層ごとに分解

$$\Theta \approx \Theta_G \oplus \Theta_H = \Theta_G \otimes \mathbf{I}_{D_H} + \mathbf{I}_G \otimes \Theta_H$$



グラフが階層構造を持つ場合に、各層のグラフ構造を効率的に推定。

Graph Product Multilayer Network型 Gauss Markov確率場の構造学習

高品佑也（早大）・井上真郷（早大）

実験:

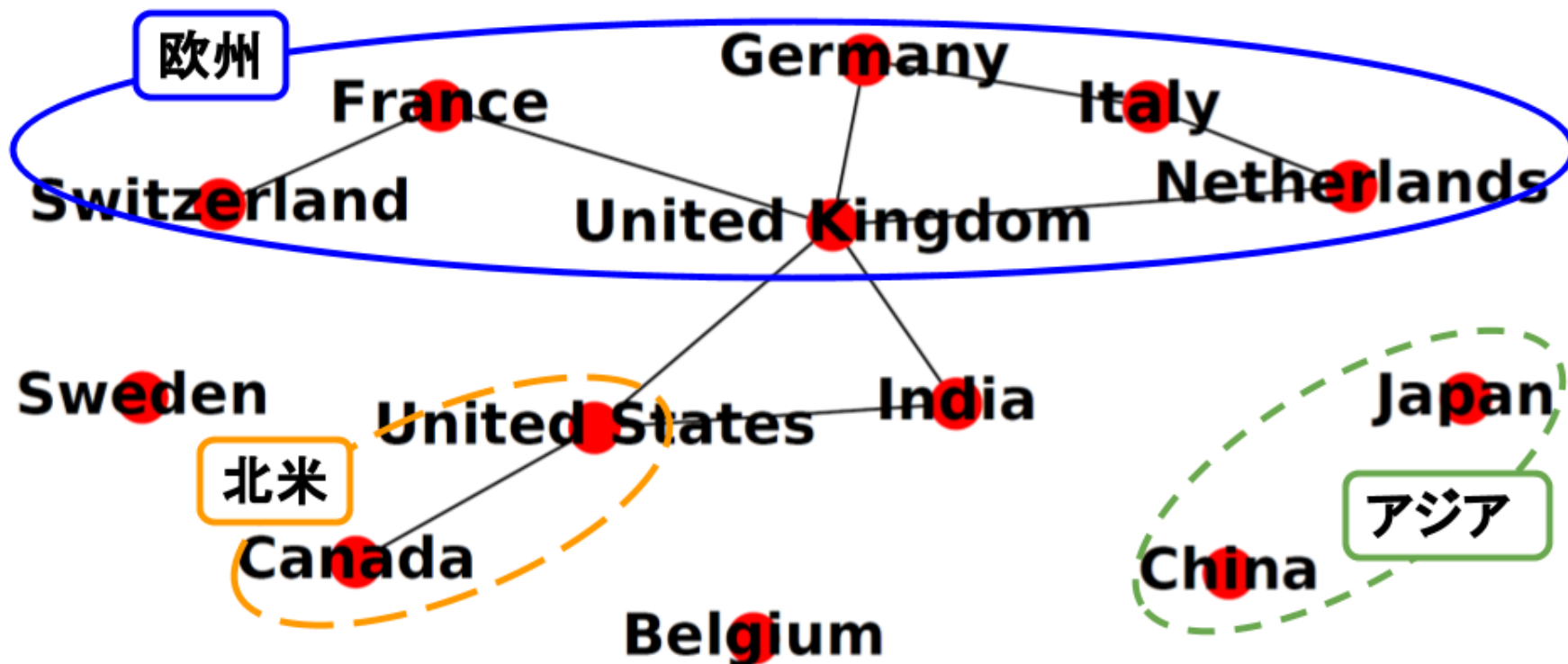
Google検索における複数国での
複数単語の検索回数のデータから
典型的な各国の関係 Θ_H を推定。

全体の
グラフ

単語ごと
のグラフ

国ごと
のグラフ

$$\Theta \approx \Theta_G \oplus \Theta_H$$

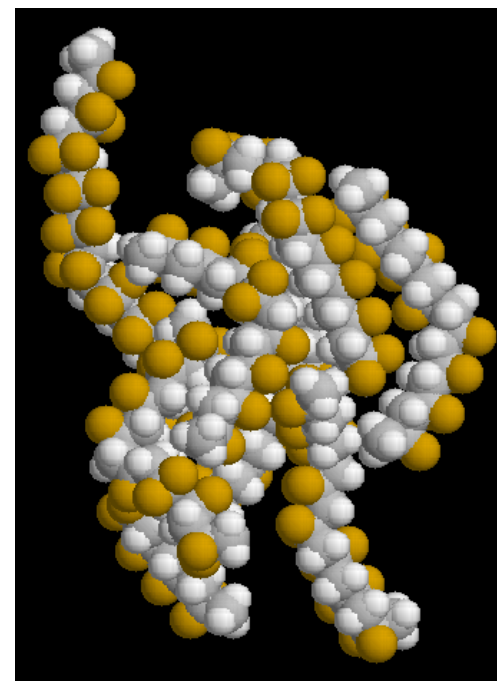


Learning Algorithm in Molecular Adaptation

Nobuyuki TAKAHASHI†

†Hokkaido Univ. of Education,

Response time of molecular systems under external field pulse was examined for conformational defect relaxation based on molecular dynamics simulation of polarization switching in comparison with machine learning algorithms.



Learning Algorithm in Molecular Adaptation

Nobuyuki TAKAHASHI†

†Hokkaido Univ. of Education, Hakodate

“temporal tuning”
Learning of short
time response by
rapid frequency
switching.

