

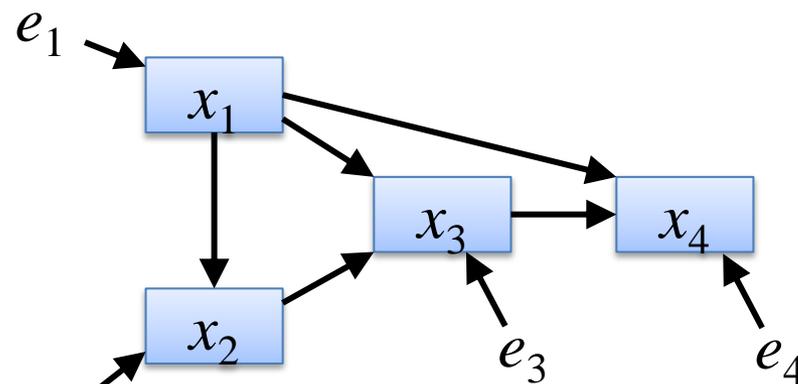
T-24 分割表の独立性に基づく二値データ生成過程の推定法

稲積 孝紀, 鷲尾 隆, 清水昌平, 鈴木 譲, 山本 章博, 河原 吉伸
(大阪大学, 京都大学)

- 二値データに対するデータ生成過程の推定

データ #	観測変数			
	x_1	x_2	x_3	x_4
#1	1	1	0	0
#2	1	0	0	1
#3	0	1	0	1
#4	1	0	1	1
#5	0	0	1	0
...

推定



$$x_1 = e_1$$

$$x_2 = x_1 \oplus e_2$$

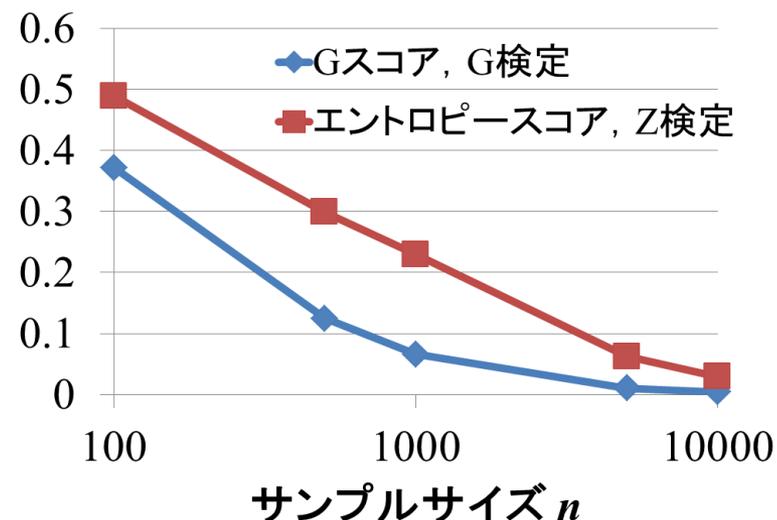
$$x_3 = (x_1 \wedge x_2) \oplus e_3$$

$$x_4 = (x_1 \vee x_3) \oplus e_4$$

- 先行研究 BExSAM [Inazumi et al., 2011, UAI]
 - 与えられたデータに対する生成過程を一意に同定できる。

- 先行研究 [Inazumi et al., 2011, UAI] における課題
 - 推定アルゴリズムに用いている指標が統計的に最適であることを保証しない.
- 本研究の目的
 - 先行研究の推定アルゴリズムに対する統計的により適切な指標を提案する.
 - G検定による独立性の検定を用いる.
- 実験結果
 - 先行研究に比べ、高い推定精度を得ることができた.

エラー率 ERs の比較



Resetting分布を用いた動的モデル選択

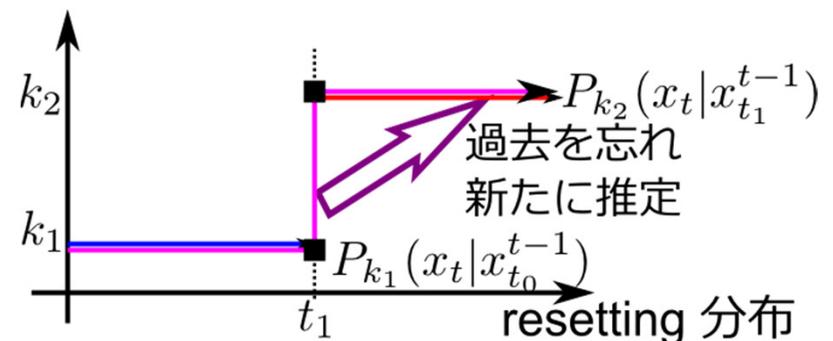
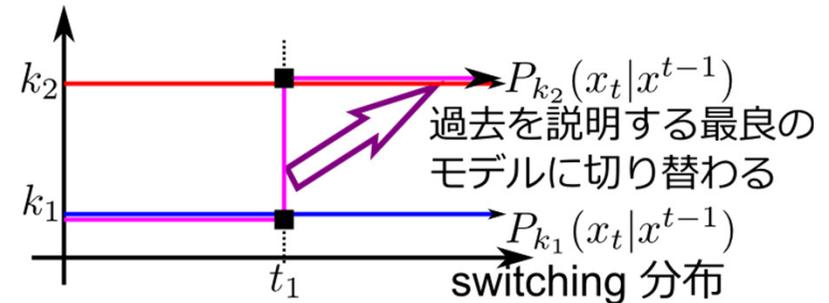
櫻井瑛一(産総研)・山西健司(東大)

動的モデル選択:
非定常な情報源の生成構造
の変化をとらえる

⇒モデル変化のとらえ方:
良いモデルに切り替わる

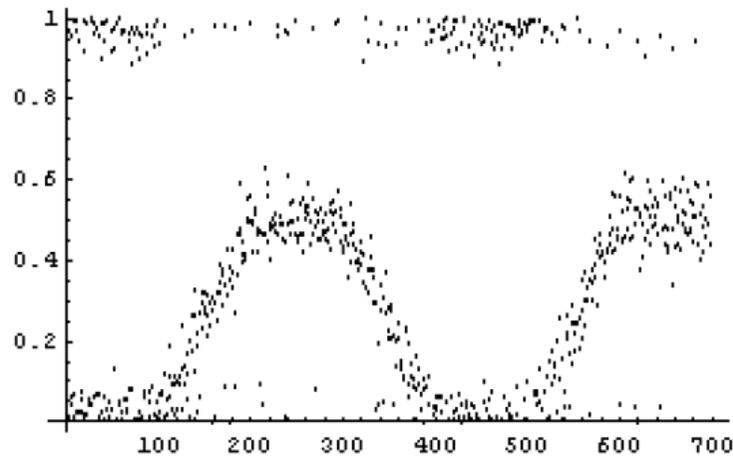


区分定常的な情報源を考え
動的モデル選択を適用

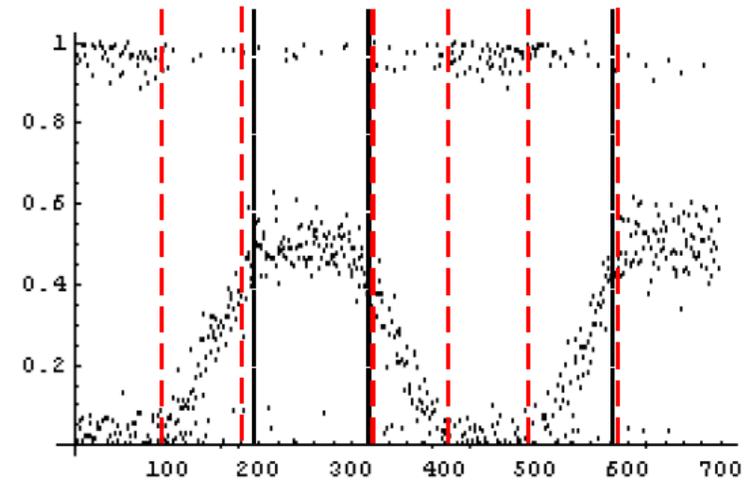


idea:記述長最小原理としてはこちらのほうがよい?

数値実験で比較



緩やかに生成構造が
切り替わる情報源からの
データ



赤:提案手法
黒:従来手法

- 変化点の検知率:提案手法>従来手法
- 変化時点の正確さは:従来手法>提案手法

ポスターではほかの場合やアルゴリズムなどを載せています

T-26 スイッチング状態空間モデルと変分ベイズ法を用いた画像の修復と領域分割

長谷川 亮太
(関西大大学院)

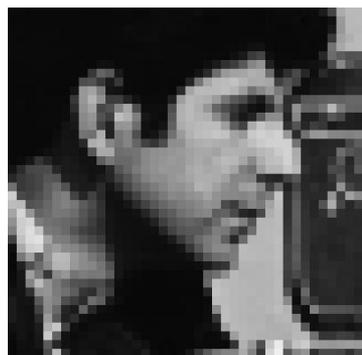
瀧山 健
(東大大学院)

岡田 真人
(東大大学院)

三好 誠司
(関西大)

➤ 変分ベイズ法を用いた画像の修復と領域分割

➤ ハイパーパラメータ推定



自然画像

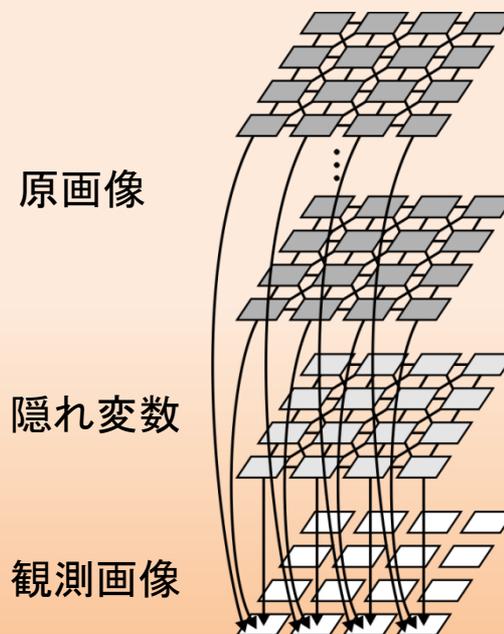
従来手法

全ての領域を単一の
スムースネスで推定

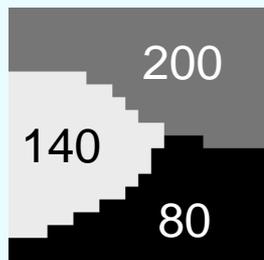
提案手法

領域ごとに異なる
スムースネスを推定

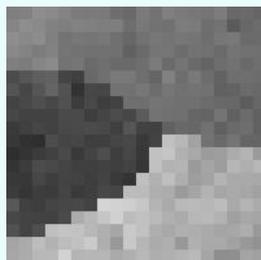
スイッチング状態空間モデル



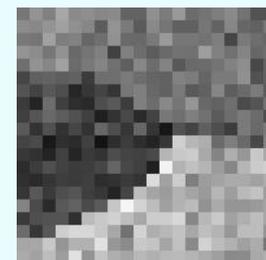
➤ 人工画像



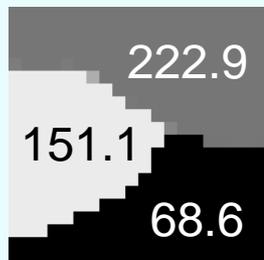
真の領域



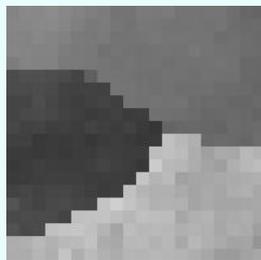
原画像



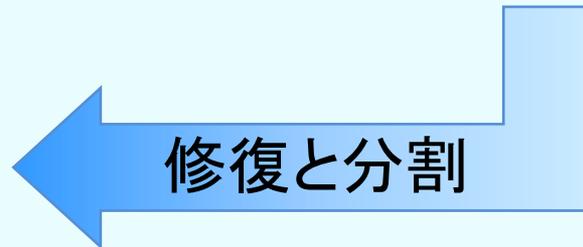
観測画像



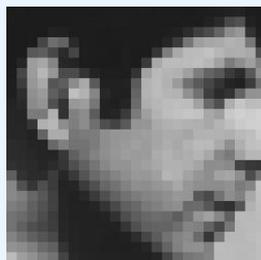
推定領域



推定画像



➤ 自然画像



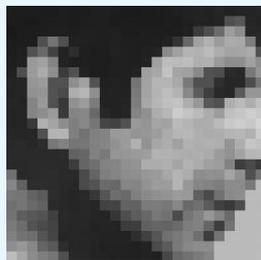
原画像



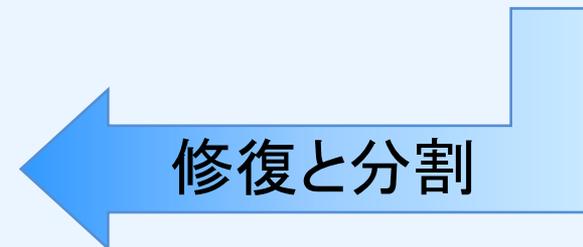
観測画像



推定領域



推定画像



T-27

変分ベイズ法とMCMCを用いた画像の修復と領域分割

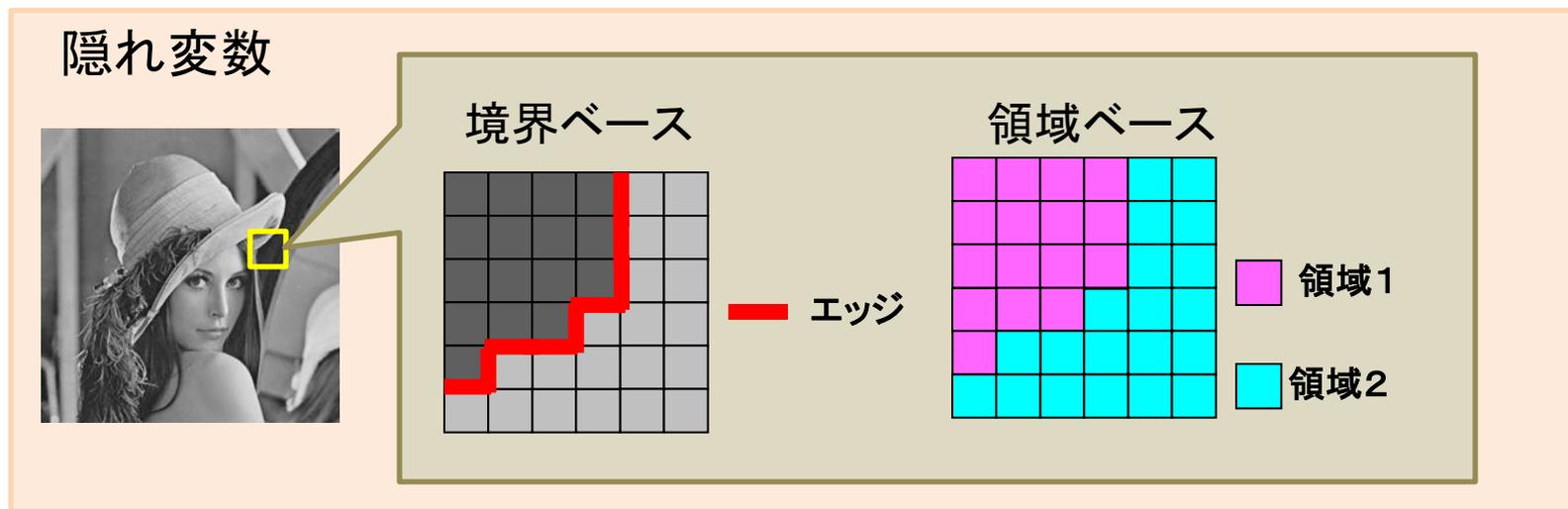
萱野 健太
(関西大大学院)

永田 賢二
(東大)

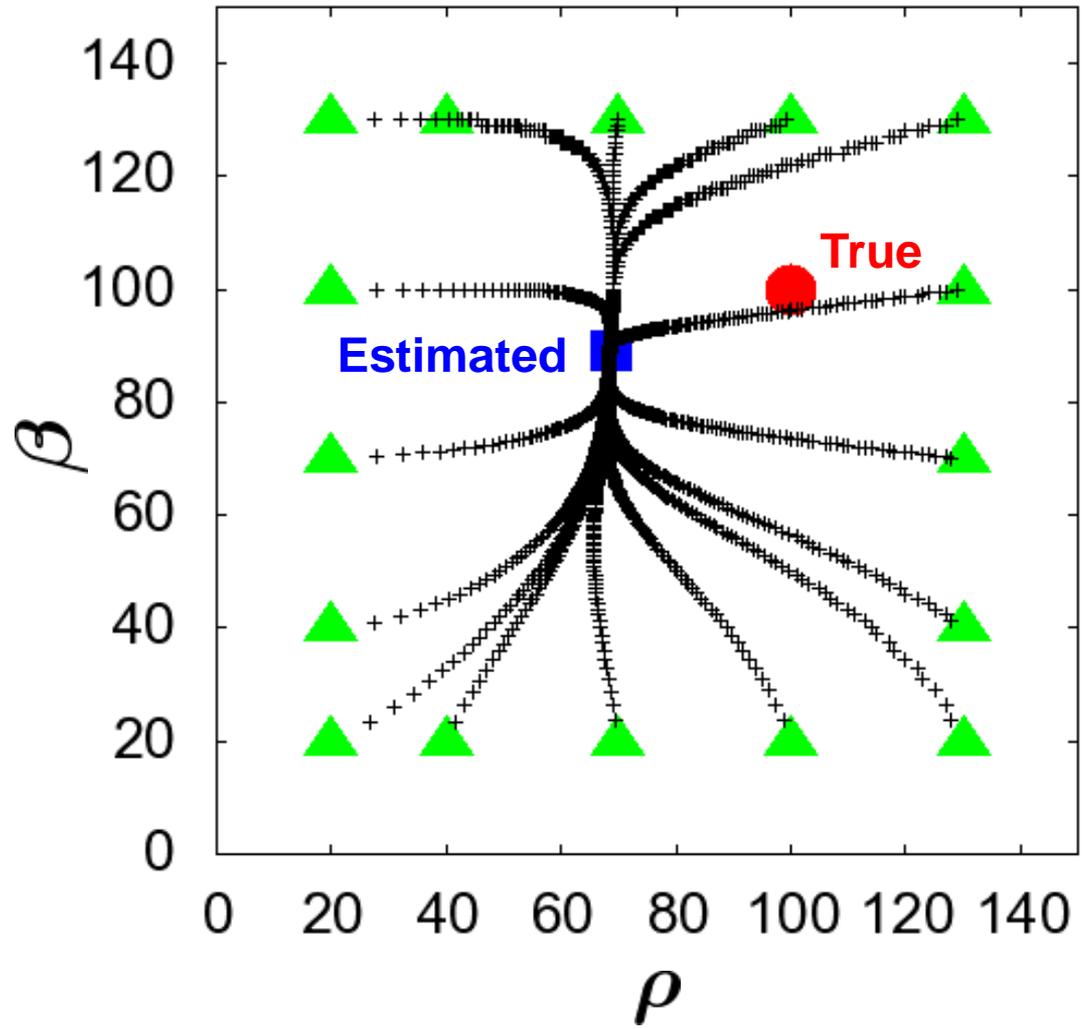
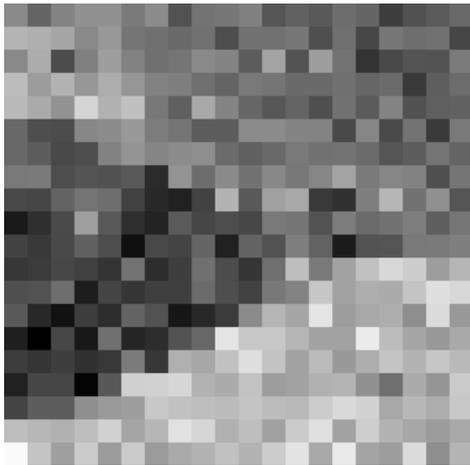
岡田 真人
(東大, 理研)

三好 誠司
(関西大)

- ◆ ベイズ推定に基づく画像処理
- ◆ エッジを表現するためには隠れ変数の導入が有効



- ◆ MCMCを用いたハイパーパラメータ推定



T-28

Dirichlet過程と変分Bayes法による交叉率推定

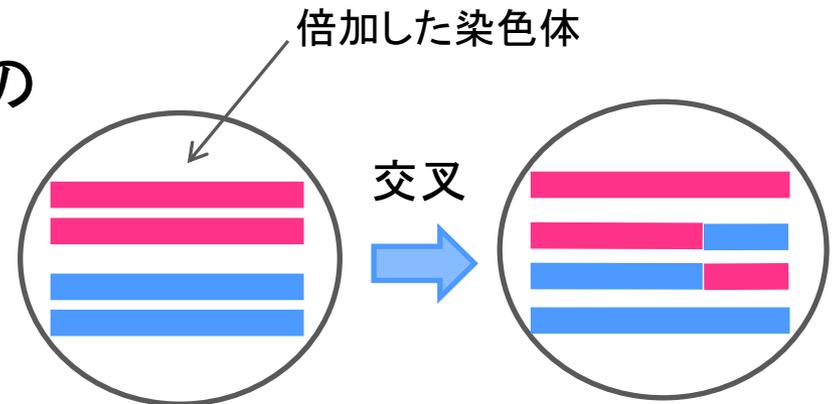
蓬田 裕菜, 村田 昇, 井上 真郷

早稲田大学大学院 先進理工学研究科

y.yomogida@ruri.waseda.jp

- 交叉

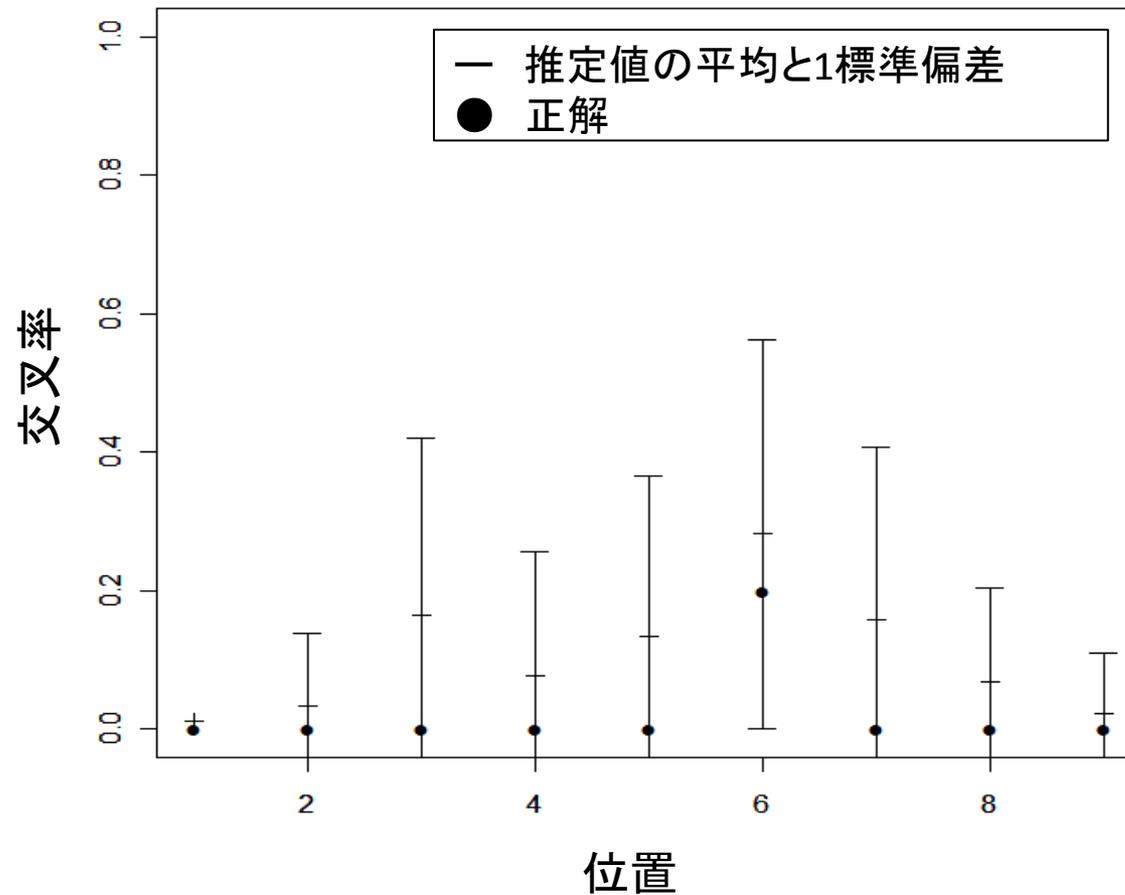
- 減数分裂時に起こる現象.
- 染色体に乗っている遺伝子の組み合わせが変化すること.



- モデル

- 隠れMarkovモデルを構築.

- 手法・・・ Dirichlet過程と変分Bayes法
 - 混合数を自動推定.
 - 解が安定せず.



Chang-Point Detection in Time-Series Data by Relative Density Ratio Estimation

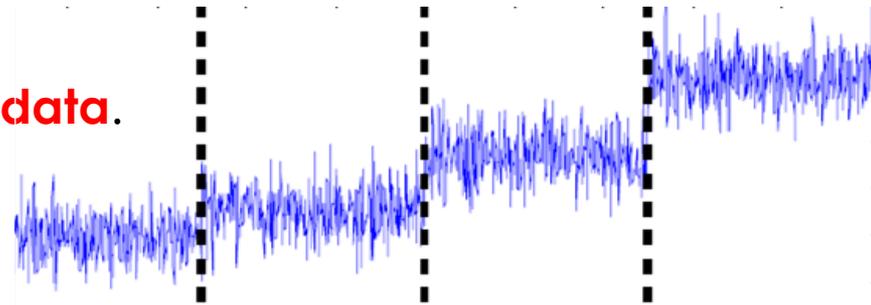
Song Liu*, Makoto Yamada*, Masashi Sugiyama*

*Tokyo Institute of Technology



Objective:

Discovering abrupt changes in **time-series data**.

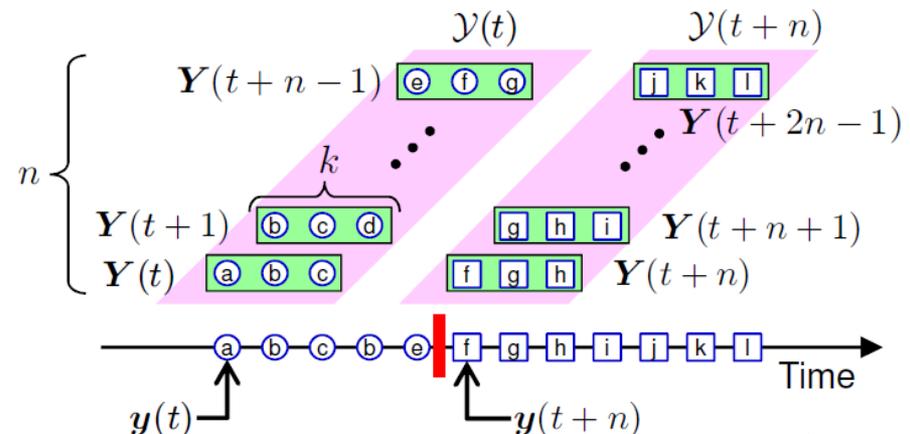


Methodology:

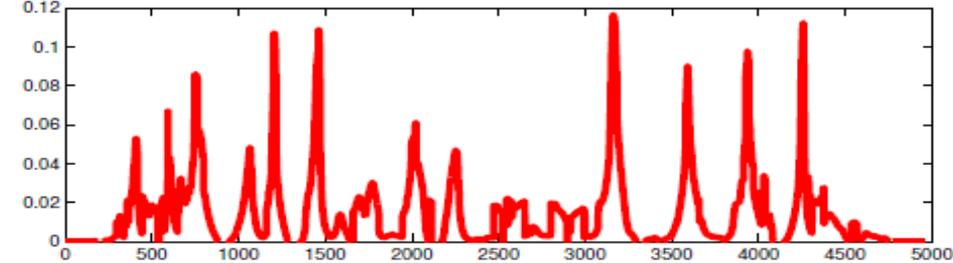
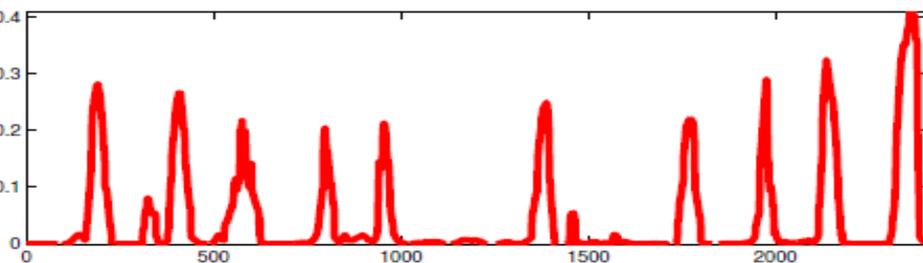
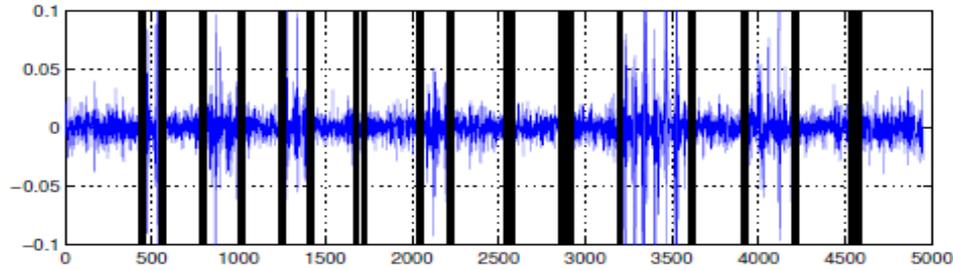
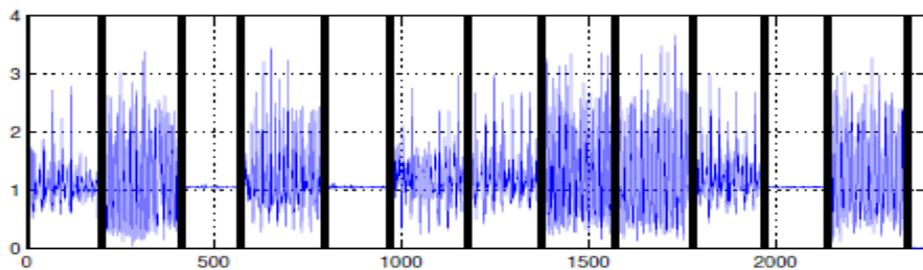
Estimating relative Pearson divergence between densities of segments by **RuLSIF**.
(Yamada et al., NIPS 2011)

Advantage:

High stability, robustness, accuracy

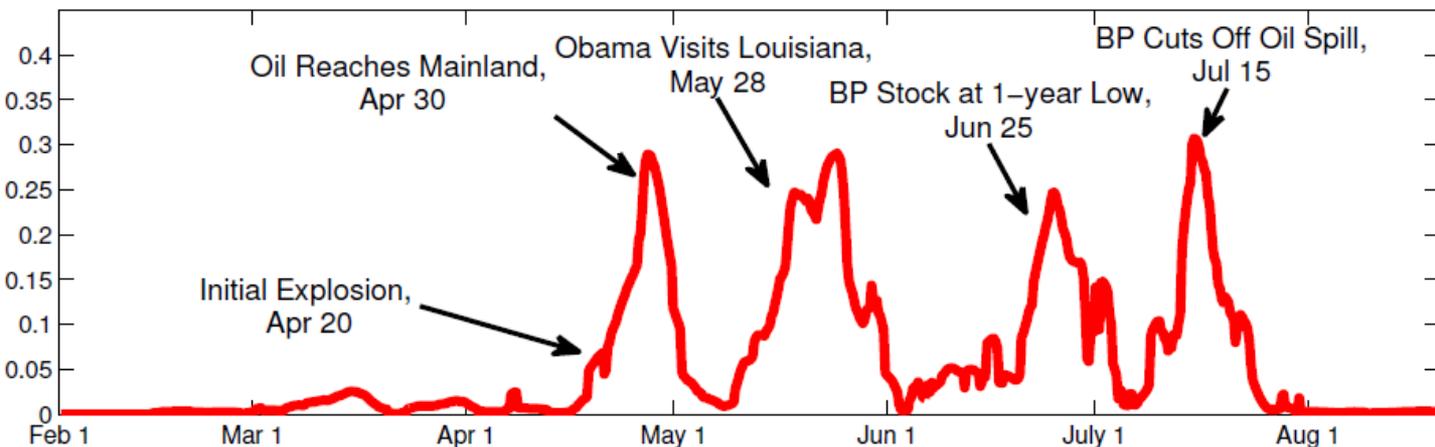
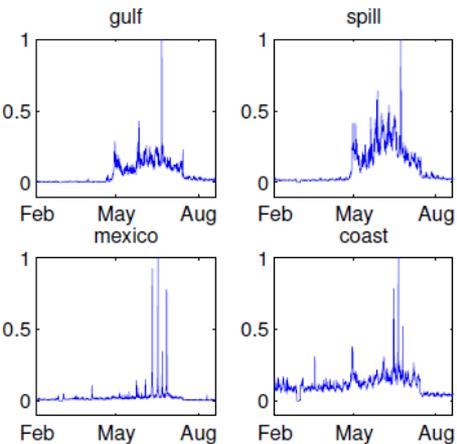


Experiments on Real World Datasets



Human Activity Segmentation

Speech Segmentation



Twitter Event Detection!

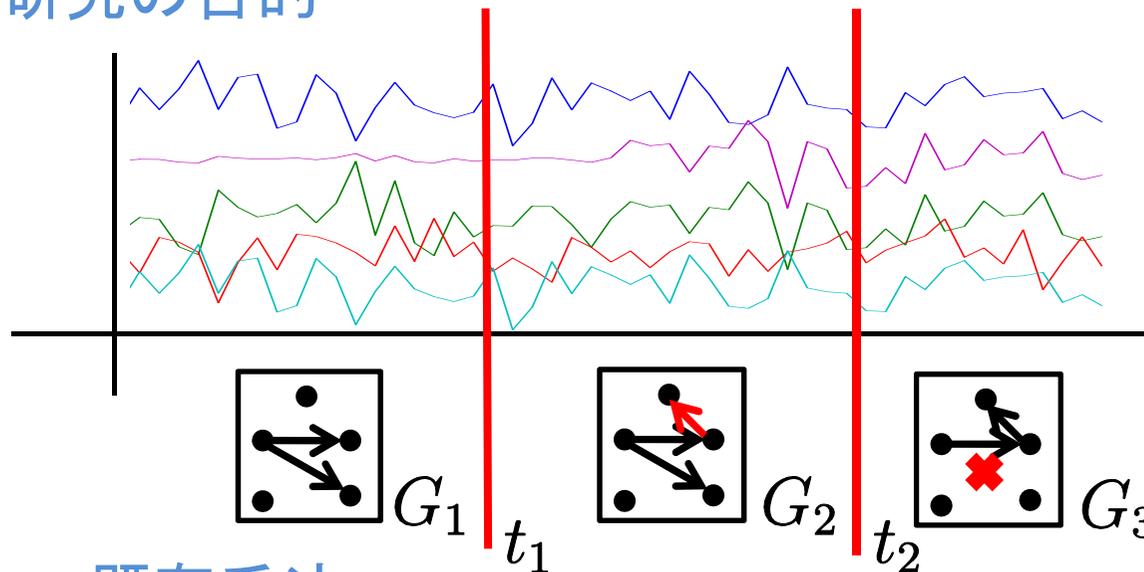


T-30: 逐次的なネットワーク構造変化検出手法と 広告効果測定への応用

早矢仕 裕, 山西 健司(東京大学)

{yu_hayashi, yamanishi}@mist.i.u-tokyo.ac.jp

研究の目的



入力: 非定常な、相関を持つ
多変量時系列データ

出力: ベイジアンネットワーク
(変数間の依存関係を表す
モデル)と変化点の組
(モデル系列)

既存手法

データ列が一括で与えられたもとでモデル系列を推定
(Xuan and Murphy, 2007 など)

提案手法

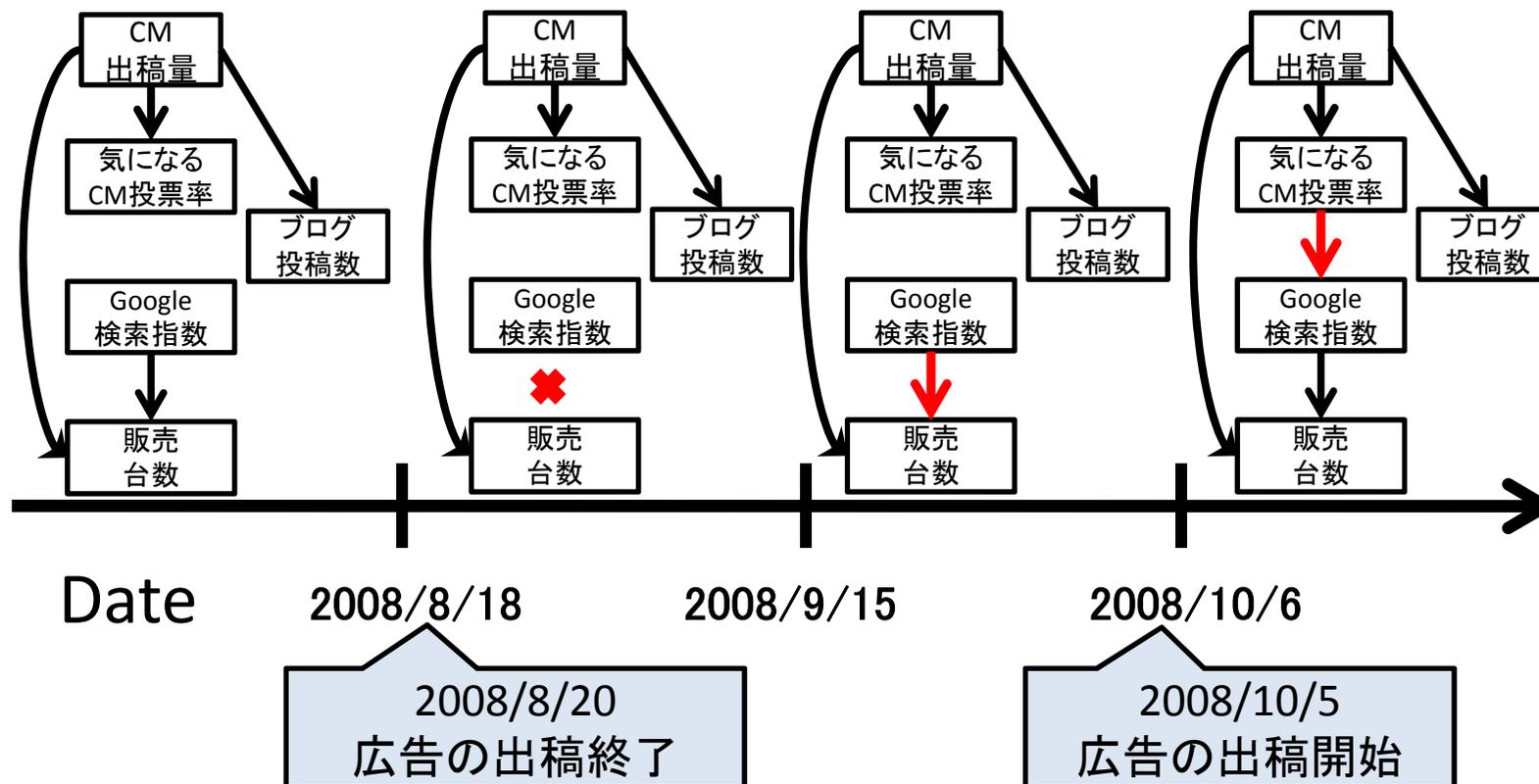
データ列が**逐次的に**与えられる状況において
モデルの変化を**リアルタイム**で検出

提案手法のポイント

- ・各時刻で一定区間データの先読みを行い、逐次的な変化検出を行う
- ・過去のモデル系列に応じて探索空間を制限する

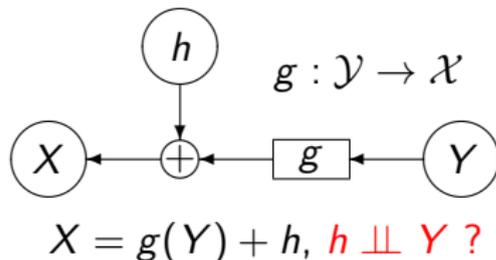
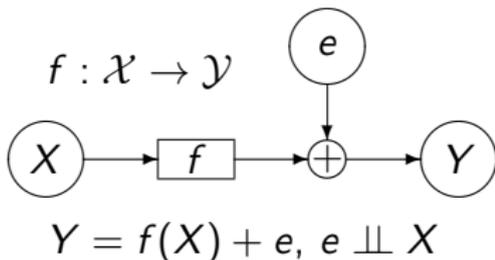
成果

- ・一括型の既存手法に比べ大きく精度を落とさず、逐次的な変化検出を行えた
- ・マーケティングデータに提案手法を適用することで、広告効果測定への応用の枠組みを示した



鈴木、清水、鷺尾: 離散データの因果の同定

X, Y : 集合 \mathcal{X}, \mathcal{Y} に値をとる確率変数



$h \perp\!\!\!\perp Y$ のとき**可逆**、 $h \perp\!\!\!\perp Y$ でないとき**因果 $X \rightarrow Y$**

- \mathcal{X}, \mathcal{Y} が連続: (X, Y) が正規分布だと可逆 (Shimizu, 2006)
- $\mathcal{X} = \mathcal{Y} = \{0, 1\}$: 鷺尾グループの BExSAM (2011)
- \mathcal{X}, \mathcal{Y} が有限: 必要十分条件が得られていない (Peters ら, 2011)

鈴木、清水、鷺尾 (阪大): 2 値から多値へ

$m = n$ 、 f : 単射、 $P(X = 0), \dots, P(X = m - 1) > 0$ のとき

定理 (可逆と同値な条件)

- $P(X = 0) = \dots = P(X = m - 1)$ 、または
- $P(e = 0) = \dots = P(e = m - 1)$ 、または
- 美しい条件 (ポスターで、 $m = 2, 3, 4$ の例を使って説明します)

応用: BN の構造推定

確率変数の集合 $A = \{X_1, \dots, X_N\}$ について、以下を繰り返す

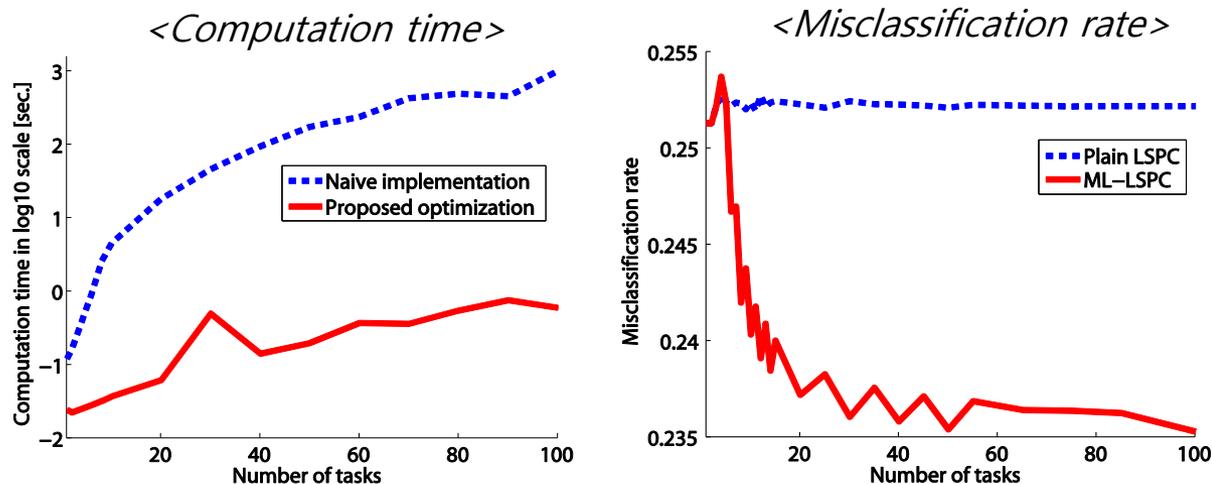
- ① 最も下流にある X_i を見出す
- ② X_i を、 X_i より上流の $\{X_j\}$ の関数 f_i と e_i の和であらわす
- ③ A から X_i を除く

[T-32] Computationally Efficient Multi-Label Classification by Least-Squares Probabilistic Classifier

Hyunha Nam • Hirotaka Hachiya • Masashi Sugiyama (Tokyo Tech)

- Multi-label classification:
 - An object belongs to multiple classes simultaneously.
 - Taking label correlations into account would be helpful.
- Least-squares probability classifier (LSPC):
Computationally efficient classification algorithm
(Sugiyama, 2010)
- This work: We extend LSPC to incorporate label correlations

- Proposed method: Multi-label LSPC
 - Efficient conjugate gradient optimization utilizing useful block structures
 (**100 times faster** than naïve implementation)
 - Label correlations help improve that accuracy
 (**1.5% less error** than plain LSPC)



- Application to real-world audio tagging
 - F-measure : **0.532 → 0.612**

T33: 局所変分法を用いた Total Variation の近似とノイズ除去

庄野 逸 

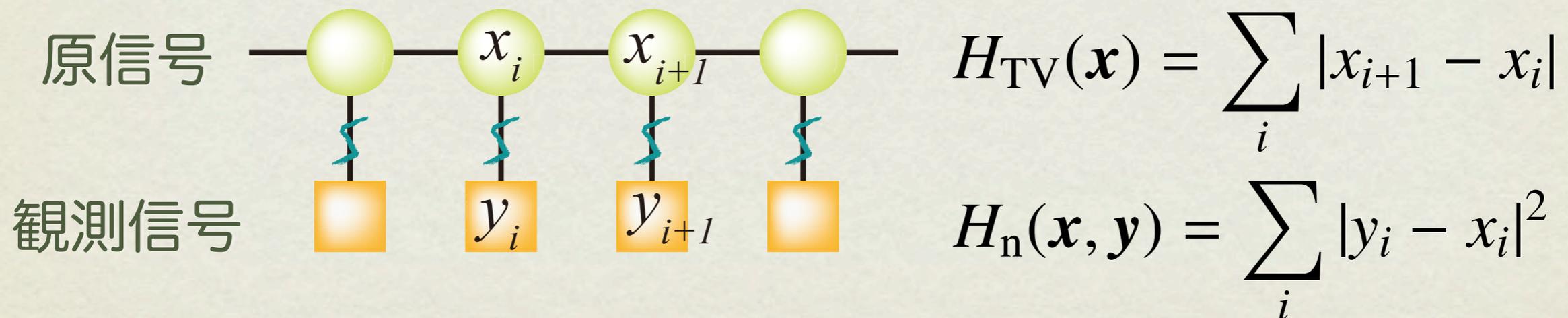
電気通信大学 大学院情報理工学研究科
shouno@uec.ac.jp

岡田 真人  東京大学
THE UNIVERSITY OF TOKYO

東京大学 大学院新領域創成科学研究科

- Total Variation (Rudin 1992)とは？

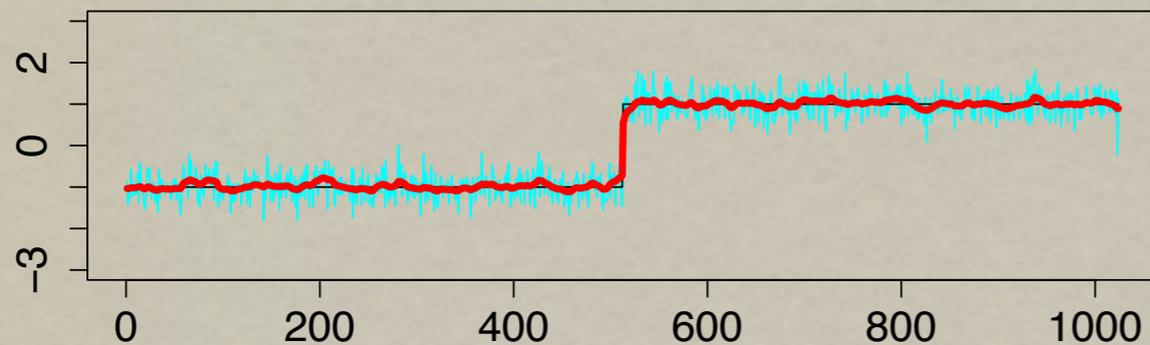
観測信号に y 含まれるノイズを除去するための拘束条件



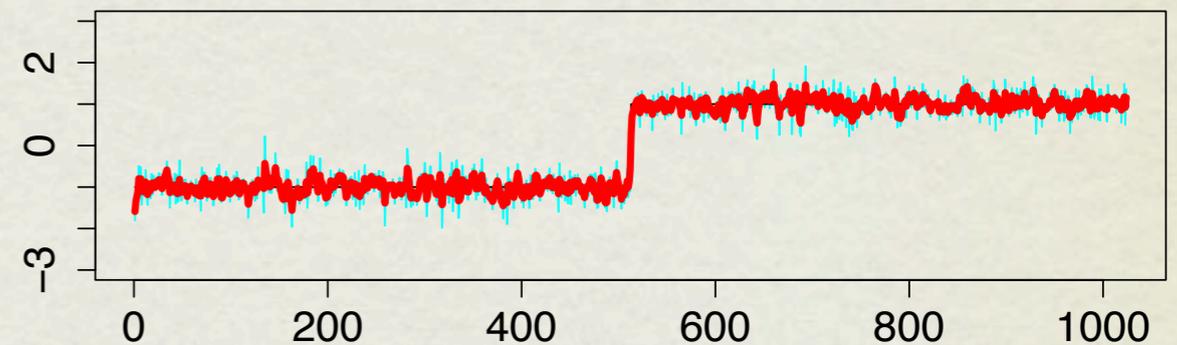
$$\min_{\mathbf{x}} H_n(\mathbf{x}, \mathbf{y}) + \alpha H_{\text{TV}}(\mathbf{x})$$

- L1-Norm: Laplace 分布の局所変分法を用いた近似
- 周辺対数尤度の最大化 by EM アルゴリズム

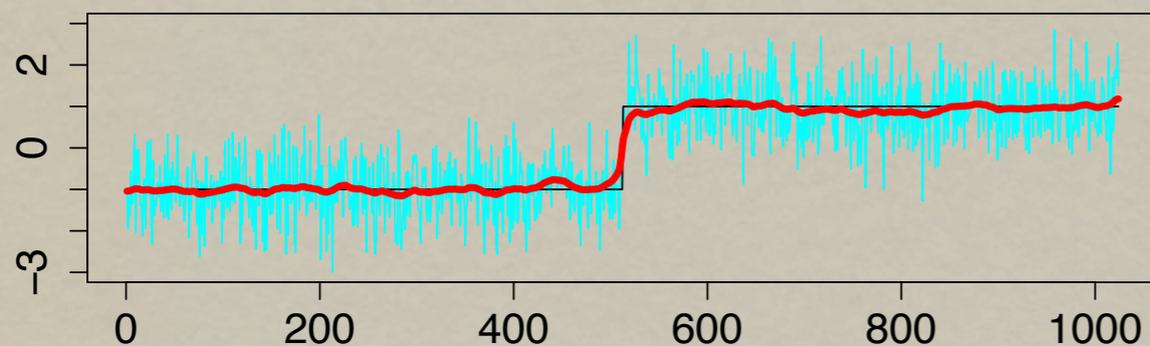
Denoised signal of $d = 0.3$ using TV



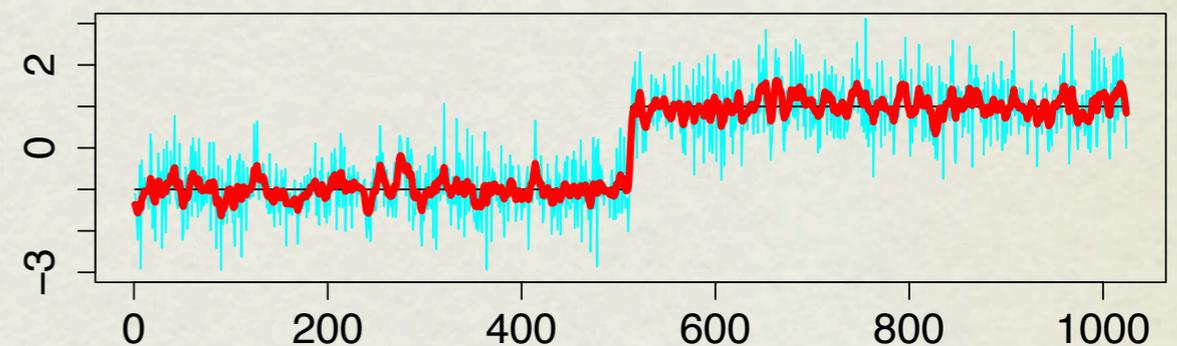
Denoised signal of $d = 0.3$ using L2



Denoised signal of $d = 0.7$ using TV



Denoised signal of $d = 0.7$ using L2



TV 近似法

L2 拘束条件

TV近似法は、効率的にノイズを除去できた

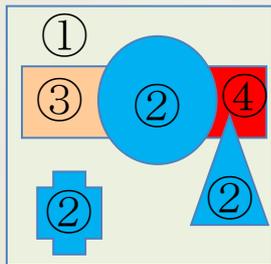
T-34 混合モデルとしての複層Gauss-Markov 確率場による 画像の修復と領域分割

勝木 孝行 井上真郷
早稲田大学
kat@toki.waseda.jp

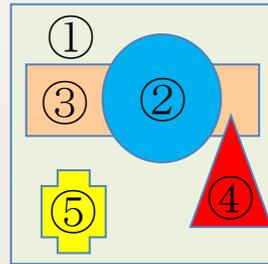
画像の修復と領域分割の同時推定



従来モデル



提案モデル



独立した物体のラベリングを行なうモデルに拡張

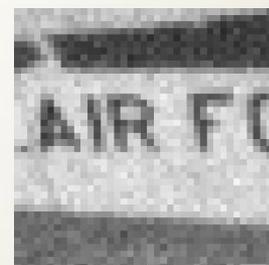
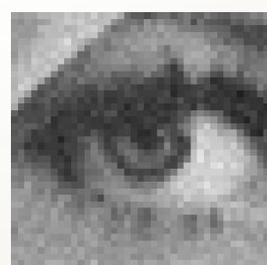
- 複層Gauss-Markov確率場
- stick-breaking-process項によるラベル数決定(モデル選択)
- フルBayesアプローチ
- 変分Bayes法
- Taylor近似

推定結果

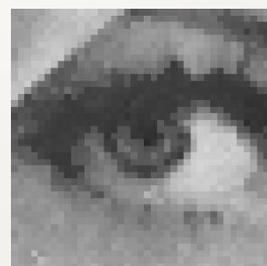
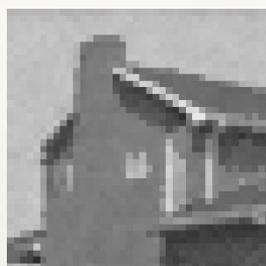
原画像



観測画像



修復結果



分割結果



①②③④⑤

①②③④⑤

①②③④⑤

①②③④

T-35 クラウドソーシングを用いた 教師付き学習の凸最適化による定式化

梶野 洸, 鹿島 久嗣 (東大)

研究成果

クラウドソーシングで得られた
信頼度の低いラベルから
直接識別器を学習する手法を提案

クラウドソーシングのタスク例: 画像が「動物」か否か

ワーカー

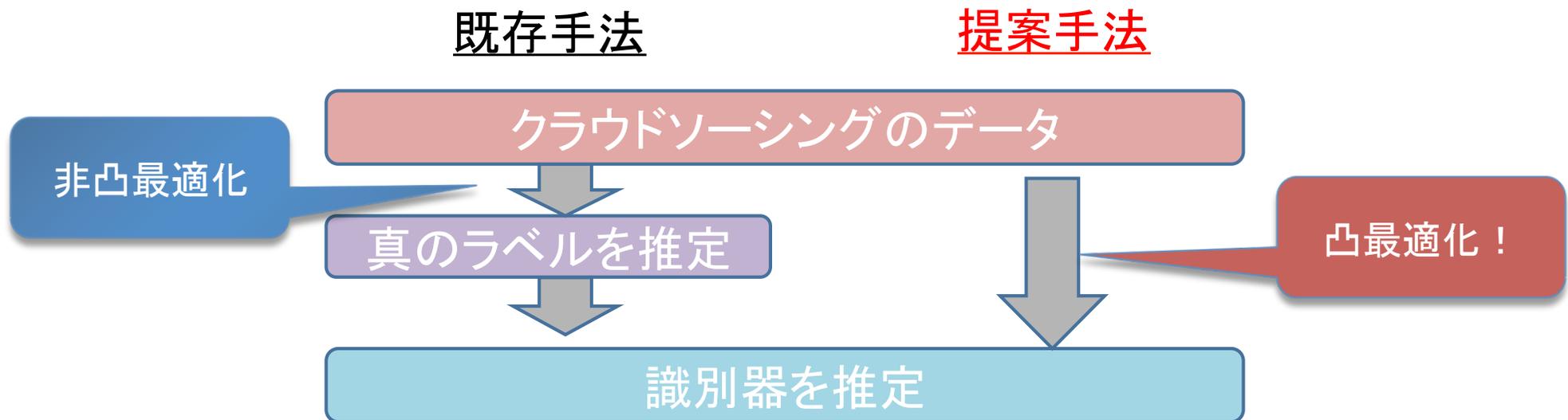
				真のラベル
	yes	yes	yes	?
	no	no	yes	?
	yes	no	yes	?

クラウドソーシングで
得られるラベルは
信頼できない

T-35 クラウドソーシングを用いた 教師付き学習の凸最適化による定式化

梶野 洸, 鹿島 久嗣 (東大)

- ▶ 既存研究とはモデルが大きく異なる
 - ▶ 凸最適化問題に帰着
 - ▶ アイデアはマルチタスク学習
 - ▶ 既存手法以上の識別性能

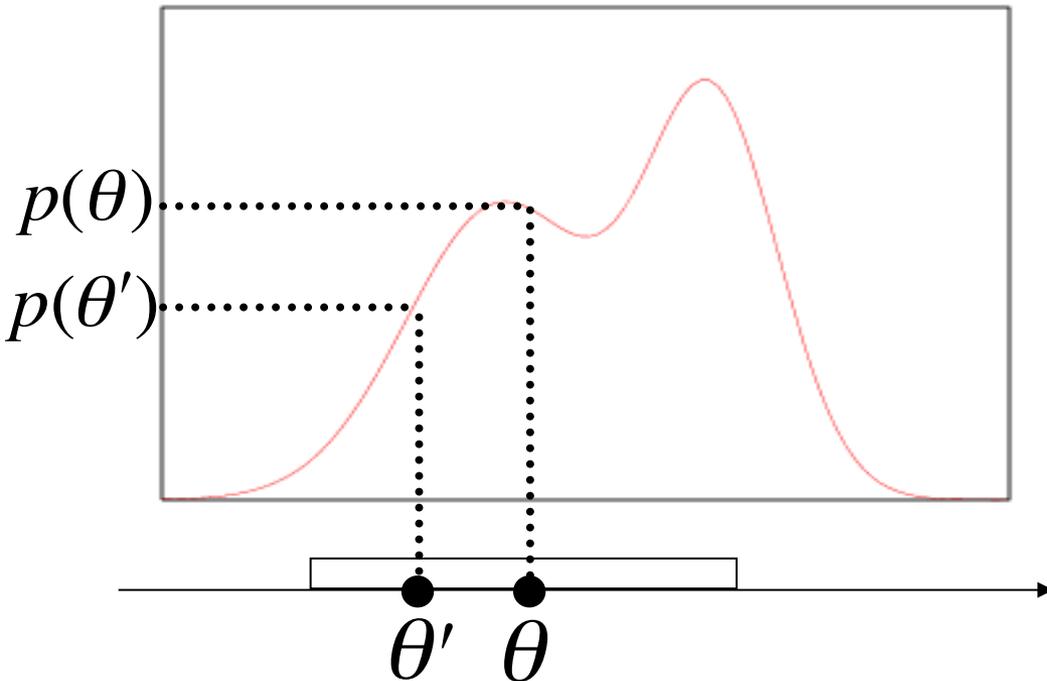


T-36 メトロポリス法における 平均採択率の挙動について

永田賢二(東京大学), 岡田真人(東京大学, 理化学研究所)

メトロポリス法

与えられた確率分布から, 経験分布(サンプリング)を生成するアルゴリズム



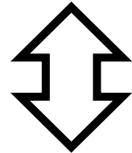
1). 候補を $\theta' = \theta + \varepsilon$ で生成

ε : 一様乱数や正規乱数など

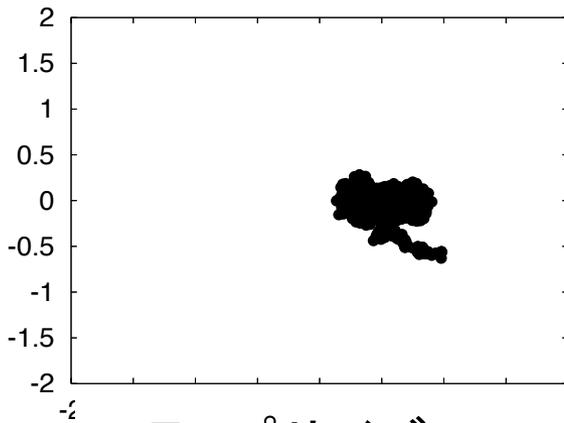
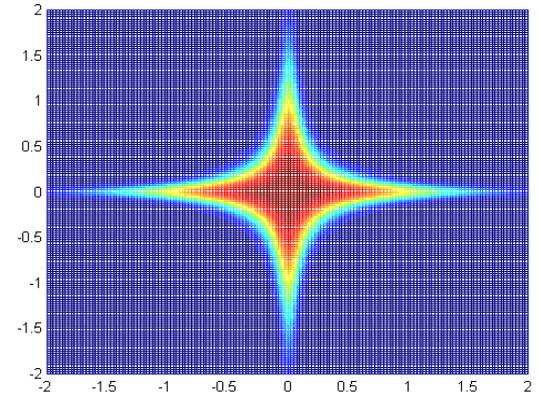
2). 比 $\frac{p(\theta')}{p(\theta)}$ に応じて確率的に採択

\mathcal{E} : 一様乱数や正規乱数など

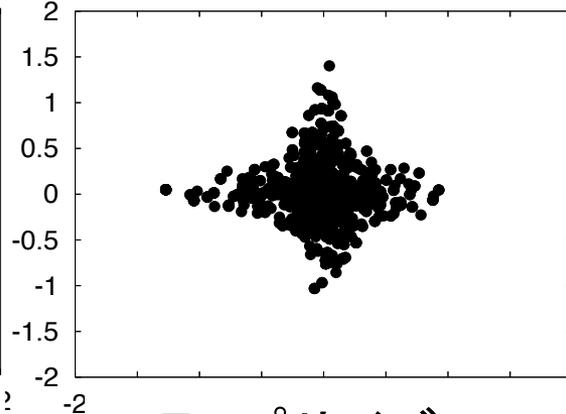
幅(ステップサイズ)の設定



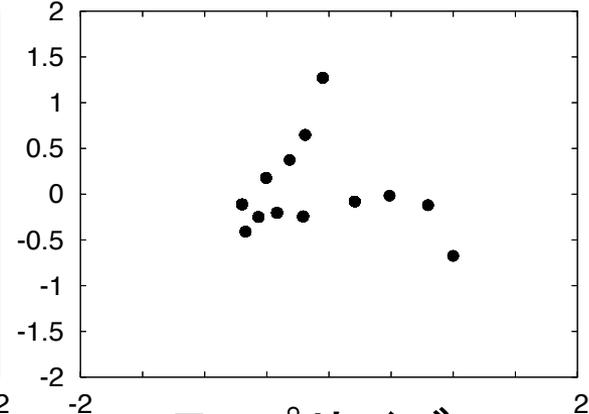
候補が採択される割合(平均採択率)



ステップサイズ:0.05



ステップサイズ:0.5



ステップサイズ:5.0

本研究では

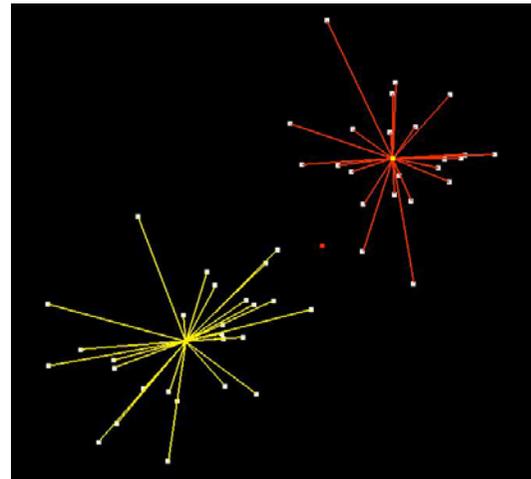
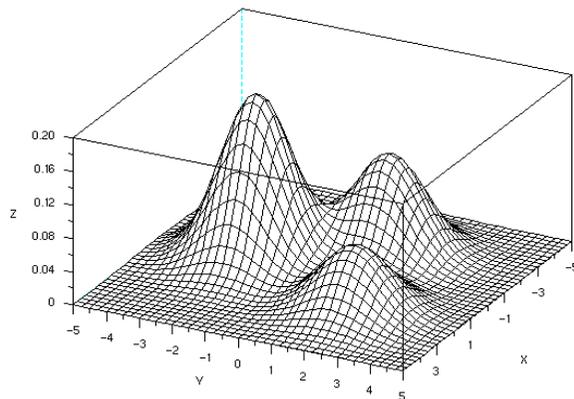
平均採択率を解析的に導出することで,

- ・目標分布とステップサイズの関係性を明らかにする.
- ・メトロポリス法や, 交換モンテカルロ法の設計指針の基盤設計

T-37:混合分布のノンパラメトリック推定と最小ダイバージェンス法

渡辺一帆(奈良先端大), 池田思朗(統計数理研)

- 混合分布の学習: クラスタリング, 密度推定
- 重み分布のノンパラメトリック推定
最尤推定量は離散分布(Lindsay)



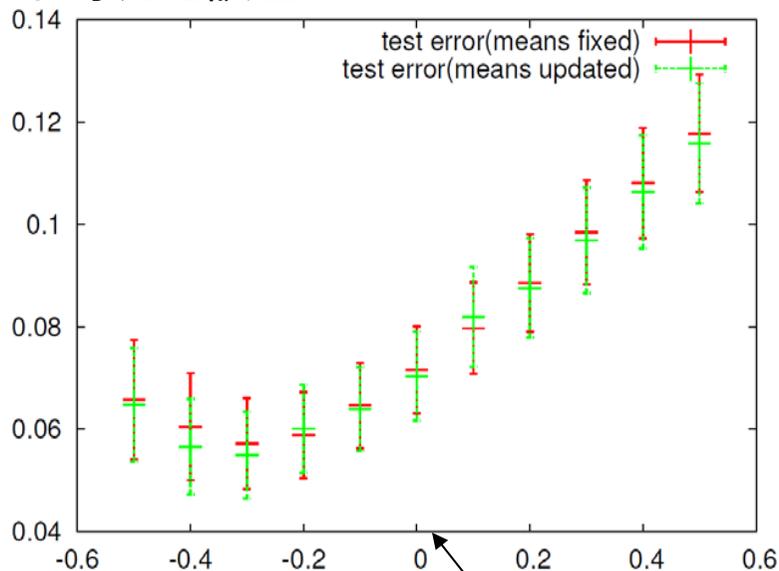
T-37:混合分布のノンパラメトリック推定と最小ダイバージェンス法

渡辺一帆(奈良先端大), 池田思朗(統計数理研)

- 混合分布 $r(x) = \int p(x|\theta)q(\theta)d\theta$ \Rightarrow $q(\theta)$ をデータから推定

Ex) $p(x|\theta)$: ガウス分布

平均汎化誤差



$$F(q) = \begin{cases} -\frac{1}{n} \sum_{i=1}^n \log r(x_i) & (\beta = 0) \\ \frac{1}{\beta} \log \left(\frac{1}{n} \sum_{i=1}^n r(x_i)^{-\beta} \right) & (\beta \neq 0) \end{cases}$$

- ダイバージェンス最小化との関係
- 過学習抑制効果 (β の範囲は?)

$\beta \rightarrow \infty$ カーネルベクトル量子化
 $\beta = 0$ 最尤推定

T-38: マルチカーネルを用いた 半教師付きドメイン適応

岡田 拓之、上原 邦昭

所属: 神戸大学大学院システム情報学研究科

E-mail: hiroyuki@ai.cs.kobe-u.ac.jp

E-mail: uehara@kobe-u.ac.jp

対象分野:

ドメイン適応、転移学習、マルチカーネル学習

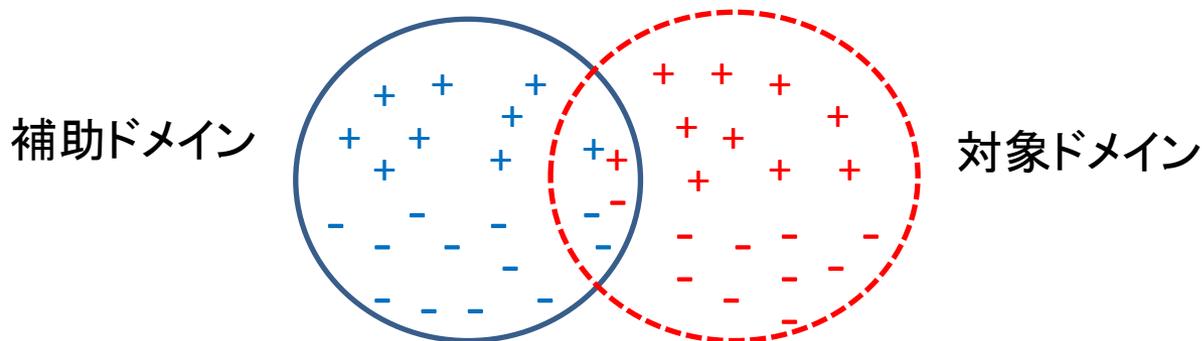
ドメイン適応とは？

- ・ドメイン適応は、転移学習の1つ
- ・ドメイン適応は大きく3つに分けられる(補助ドメインにはラベルありデータがある)

教師ありドメイン適応.....対象ドメインにラベルありデータがたくさんある

半教師付きドメイン適応...対象ドメインにラベルありデータが少量でラベルなしデータがたくさんある

教師なしドメイン適応.....対象ドメインにラベルなしデータがたくさんある



従来手法の問題点

- ・従来のドメイン適応で、データを高次元空間へ写像して問題を解く方法があった
- ・高次元空間への写像関数は実験的に決められる

本研究のアプローチ

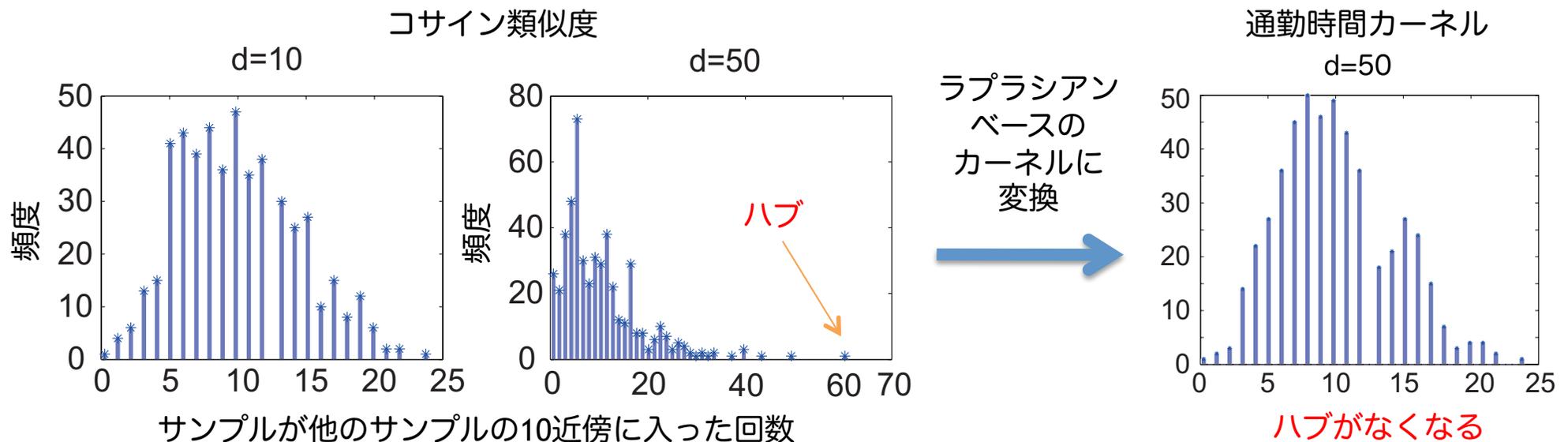
- ・マルチカーネル学習を導入して、データに適した高次元空間への写像関数を得る

T-39, 「ハブの出現しやすさ」から見たラプラシアンベースカーネル

鈴木郁美(奈良先端大), 原一夫(遺伝研), 新保仁(奈良先端大), 松本裕治(奈良先端大)

- ① ラプラシアンベースのカーネル：サンプル間の類似度を測る類似度尺度
- ② ハブ：ある類似度尺度でサンプル間の類似度を測り， k 近傍リストを作成したとき，多くの（他の）サンプルの k 近傍に出現するサンプルのこと。

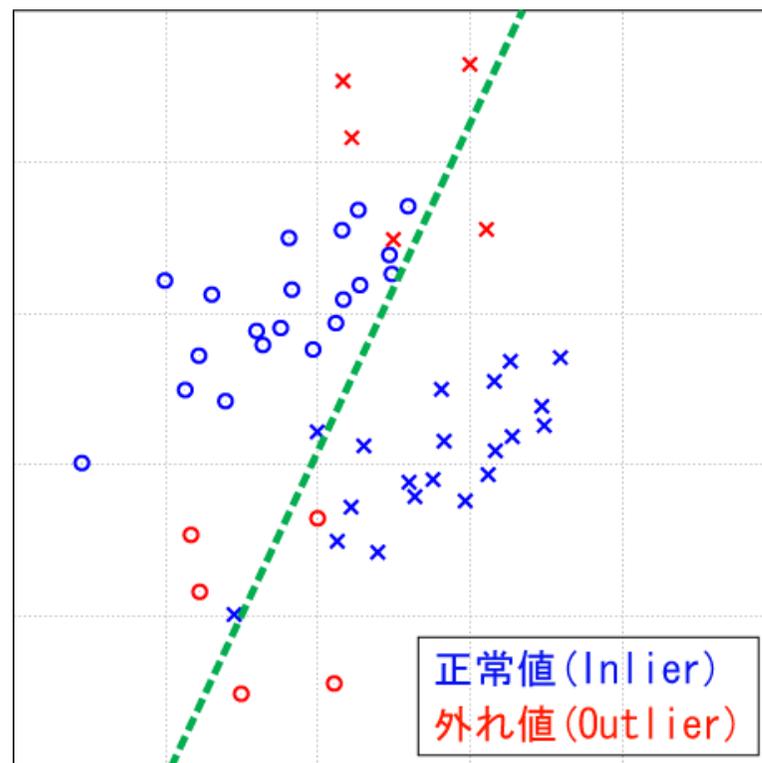
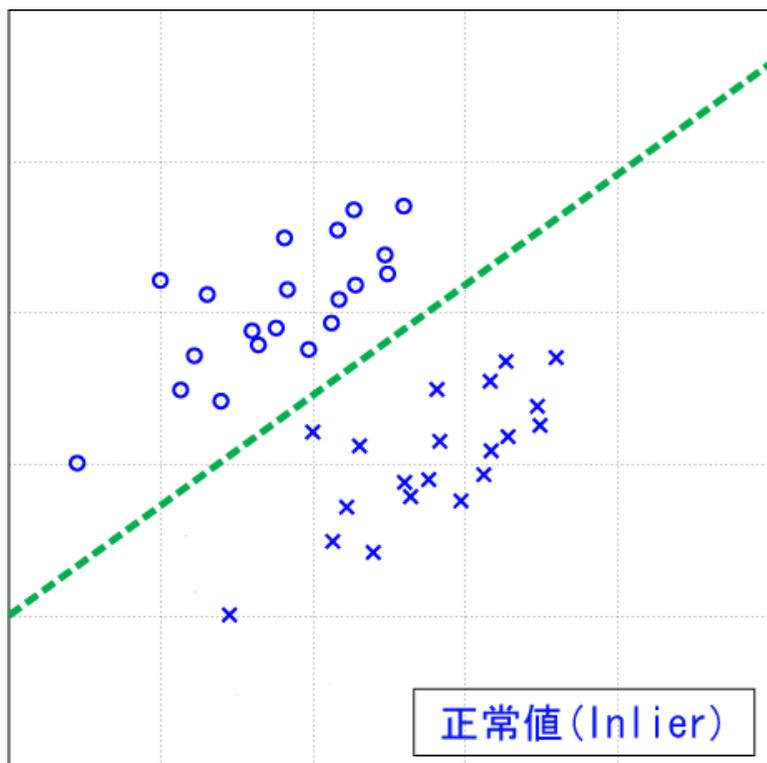
高次元データではハブが出やすい [Radovanovic et al. 2010].



本研究では，ラプラシアンベースのカーネルが，
どうしてハブを作りにくいのか，理由について考えて見ました。

T-40: パラメトリック計画法を用いた 分類と回帰における外れ値検出とロバスト学習

竹内一郎 (名古屋工業大学)



2 クラス線形分類問題における外れ値除去

T-40: パラメトリック計画法を用いた 分類と回帰における外れ値検出とロバスト学習

竹内一郎 (名古屋工業大学)

- 重み付きデータに対する最適化問題を考える

$$\begin{aligned} \min_f \quad & \sum_{i \in \text{Outlier}} w_i \text{loss}(y_i, f(x_i)) + \sum_{i \in \text{Inlier}} w_i \text{loss}(y_i, f(x_i)) \\ \text{s.t.} \quad & \text{residual}(y_i, f(x_i)) \geq \text{threshold}, i \in \text{Outlier} \\ & \text{residual}(y_i, f(x_i)) \leq \text{threshold}, i \in \text{Inlier} \end{aligned}$$

- 重みの変化に応じた最適解パスを追跡する

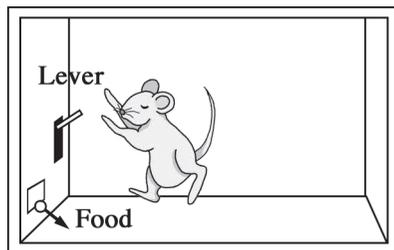
$$\begin{aligned} w_i : & \implies 0.0, i \in \text{Outlier} \\ w_i : & \implies 1.0, i \in \text{Inlier} \end{aligned}$$

T-41 学習機械間の動作により言語を生成する

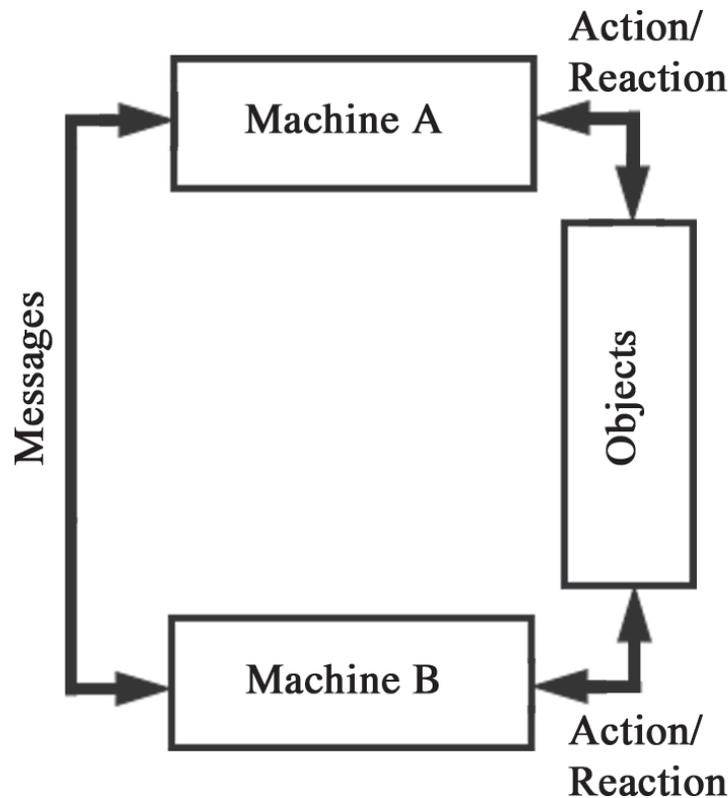
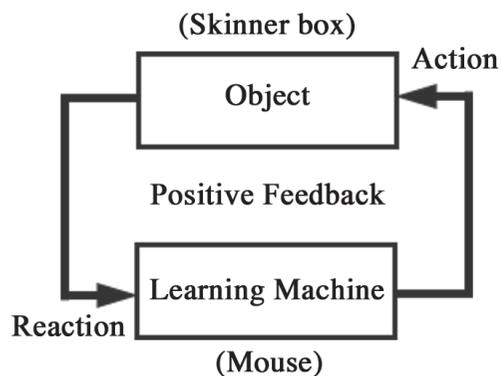
～ 自然言語発生のモデル化 ～

柳川誠介 (オプティッド)

s_yanagawa@jcom.home.ne.jp

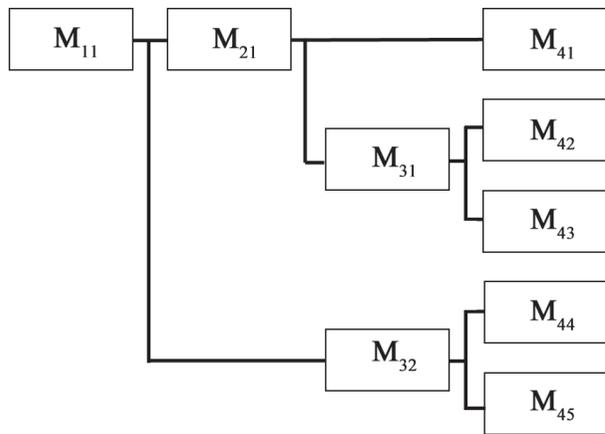


動物の学習行動を正帰還の条件発見の過程と捉える。

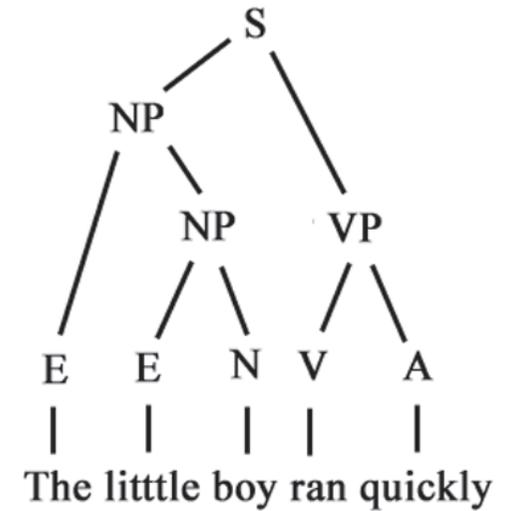


同等な機能を持つ
2個の学習機械を考
え、入出力は有限長
の記号列(文)とする。

共有文の有無で学
習機械の集合の同値
類を定める。

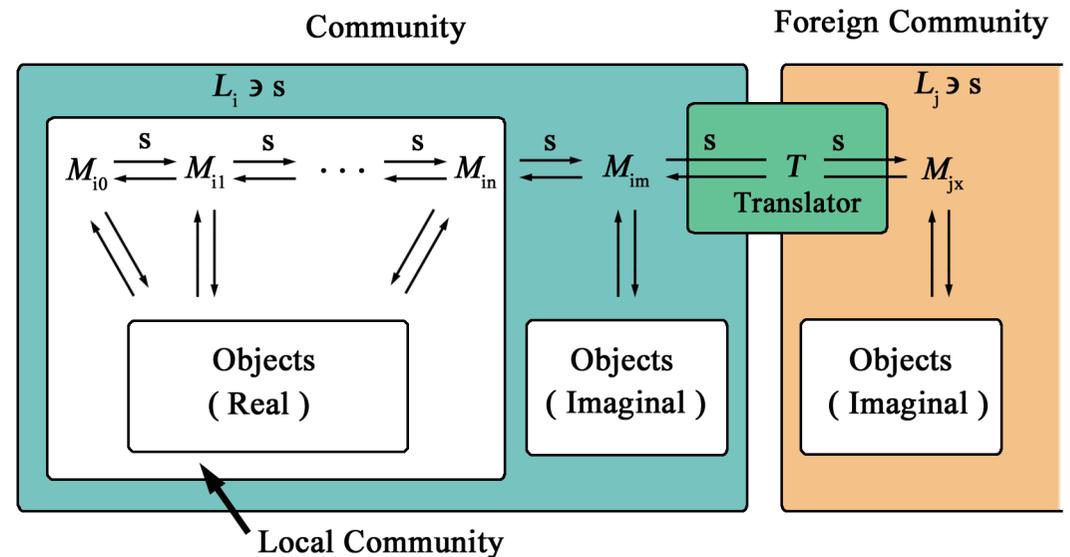
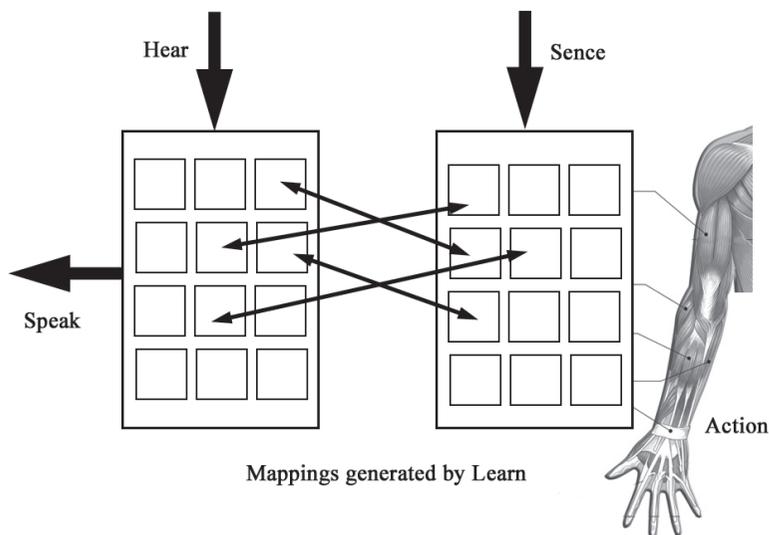


学習機械を記号列変換器と考え、階層的に接続すれば文法構造を持った文の処理が出来る



Object との Action/Reaction を行う部分と文の送受を行う部分に分割。腕の動きにも文法構造が内在する。

ヒトが使ってきた言語と関わることは可能か？



T-42 予測尤度最大化に基づく誤分類サンプルの検出

藤野昭典, 岩田具治, 永田昌明

NTT コミュニケーション科学基礎研究所

課題: 誤分類サンプル検出のための確率モデル設計

特徴ベクトル: x_n

クラス: $y_n \in \{1, \dots, k, \dots, K\}$

確率モデル

y_n に所属する確率の推定

$\hat{P}(y_n | x_n)$

誤分類サンプルが混在するラベルありデータ集合 $D = \{(x_n, y_n)\}_{n=1}^N$ のみで学習

誤分類サンプルの悪影響を低減させるためのアプローチ

1. 正誤を表す潜在変数 $z \in \{z_1, z_0\}$ を用いた生成モデル

$$p(x, y; \Theta) = \sum_{z \in \{z_1, z_0\}} p(x, y | z; \Theta) P(z)$$

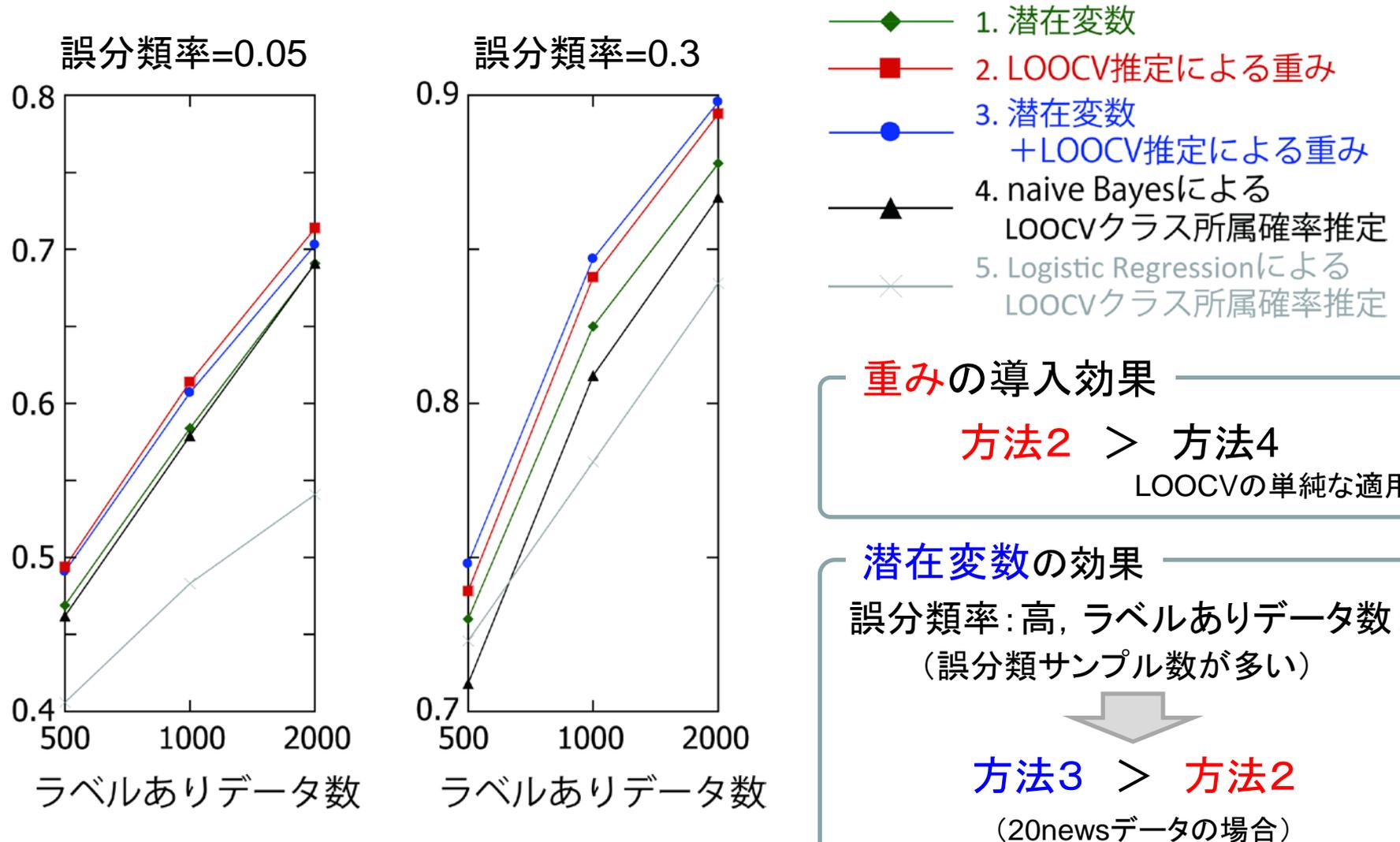
2. LOOCV 尤度最大化に基づいて推定する重みを導入したパラメータ学習

$$J(\{\theta_{y,-n}\}_{y,n}) \equiv J(\mathbf{w}) = \sum_{n=1}^N \log p(x_n | y_n; \hat{\theta}_{y_n, -n}(\mathbf{w})) P(y_n)$$

$$\hat{\theta}_{y,-n}(\mathbf{w}_1) \equiv \operatorname{argmax}_{\theta_y} \left\{ \sum_{n' \neq n} w_{n'} I_{y_{n'}}(y) \log p(x_{n'} | y; \theta_y) + \log p(\theta_y; \eta) \right\}$$

テキストデータによる評価実験

誤分類サンプルの順位付けの平均適合率 (20newsデータ)



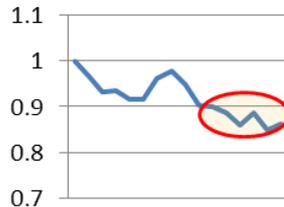
T-43

2項分布に基づく基底関数による混合回帰モデル

中瀬 弦, 井上 真郷
早稲田大学

gen-endeavor@akane.waseda.jp

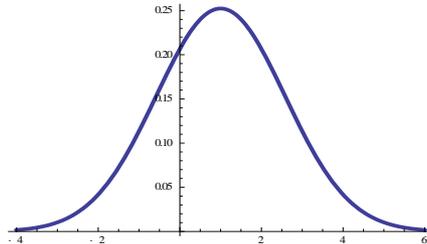
Markov性に基づき, 連続値データの分布予測を行いたい



この波形から次のデータの
予測分布を作成する

従来手法

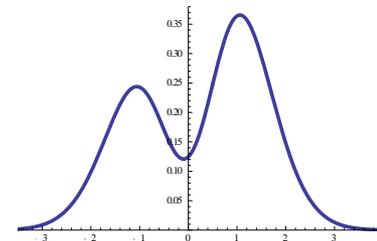
Gaussian Process (GP)



Gauss分布に従う分布推定

提案手法

基底関数の混合による予想

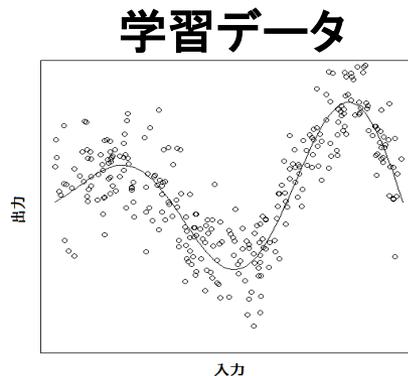


分布形に依らない推定

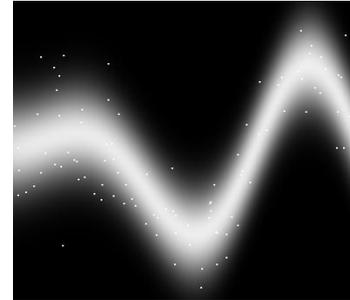
入力によって出力が確率的に決定する(多峰性)データの予測が可能

分布予測例

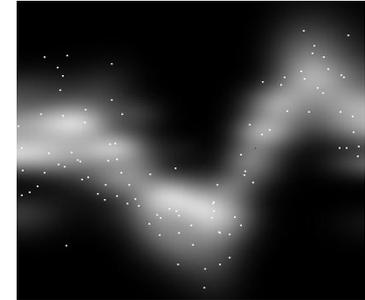
単峰性データ



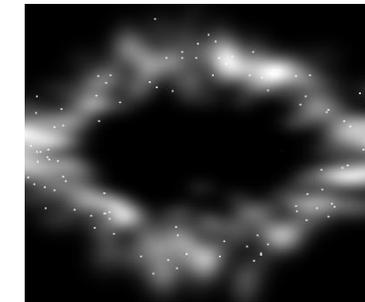
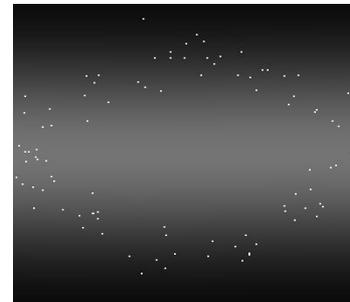
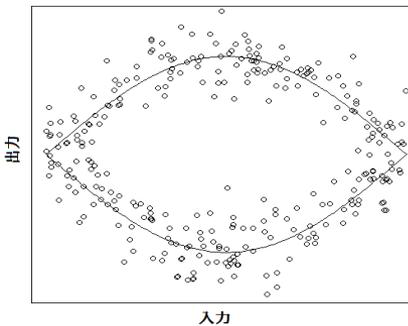
GP



提案手法

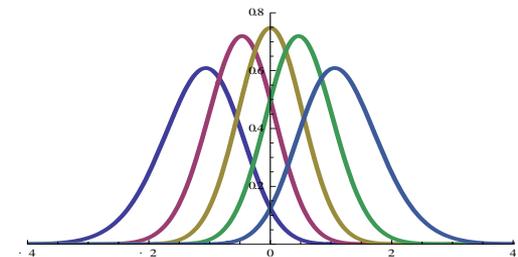


多峰性データ



特徴

- 離散値データの圧縮手法と同様のモデルによる学習
- 2項事後分布による自然な基底分布



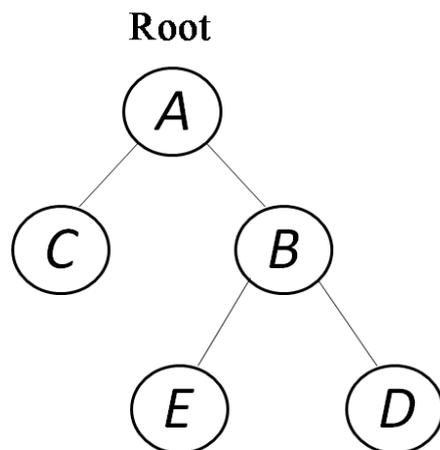
Markov性に基づく人工データにおいても有効な予測が可能

T-44

部分パスに基づいた線形時間木カーネル

木村大翼(東大), 鹿島久嗣(東大)

- ▶ 木構造を扱う学習問題に対するカーネル法の適用を考える
 - ▶ 汎用性をもち、かつ線形時間の高速な木カーネル
- ▶ 提案法ではカーネル関数を定義する特徴として、縦方向の部分パス(木の根から各葉までのパスの一部)を用いる



部分パス集合
による特徴抽出



全部分パス集合

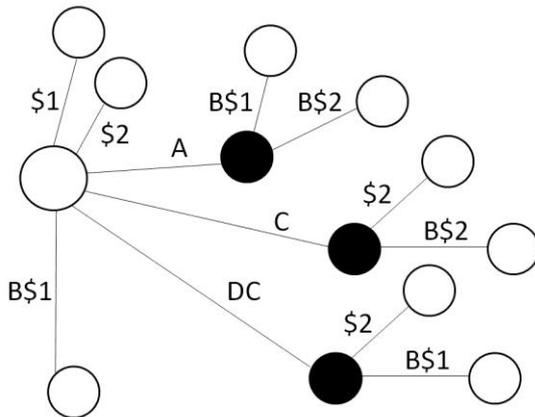
$\{A\}, \{B\}, \{C\}, \{D\}, \{E\},$
 $\{AC\}, \{AB\}, \{BE\}, \{BD\},$
 $\{ABD\}, \{ABE\}$

- ▶ カーネル関数の計算に関して線形時間アルゴリズムを構築

木構造に対する拡張接尾辞配列を用いた線形時間アルゴリズム

- ▶ 最も単純には、木構造に対する**接尾辞木**を用いることで線形時間に
 - ▶ しかし、接尾辞木は大量のメモリを消費するので好ましくない
- ▶ 代わりに**拡張接尾辞配列** (接尾辞配列 + 高さ配列)を用いたより省メモリなアルゴリズムを構築
 - ▶ 文字列カーネルの場合の線形時間アルゴリズムの拡張

× 接尾辞木



○ 拡張接尾辞配列

i	SA	LCP	
1	1	0	$\$1$
2	6	0	$\$2$
3	3	1	$AB\$1$
4	8	0	$AC\$2$
5	2	0	$B\$1$
6	7	1	$C\$2$
7	4	0	$CB\$1$
8	9	2	$DC\$2$
9	5	-1	$DCB\$1$

- ▶ 実験的に省メモリかつ計算時間は線形にスケール

階層Dirichlet過程を用いた3次元計測データのクラスタリングと行動計画

下斗米貴之・岡田浩之・大森隆司(玉川大)

背景: 自律ロボットにとって、各種センサを通じた実世界の認識が重要
地図の構築(SLAM) (e.g. S.Thrun et al. 2005), 物体概念の構築(Nakamura, 2011)

目的:

確率的な推定によるロボットの外界の認識機構の構築と開発

システム:

入力: 赤外線3次元計測データ
予測システム

モデル:

階層型混合ガウス分布

$$p(x_{ti}|a, j) = \frac{1}{2\pi|S_x|} \exp \left[-\frac{1}{2} \sum_i (x_{ti} - a_{tj})^T S_x^{-2} (x_{ti} - a_{tj}) \right]$$

$$p(v_i) = \frac{1}{2\pi|S_v|} \exp \left[-\frac{1}{2} v_i^T S_v^{-2} v_i \right]$$



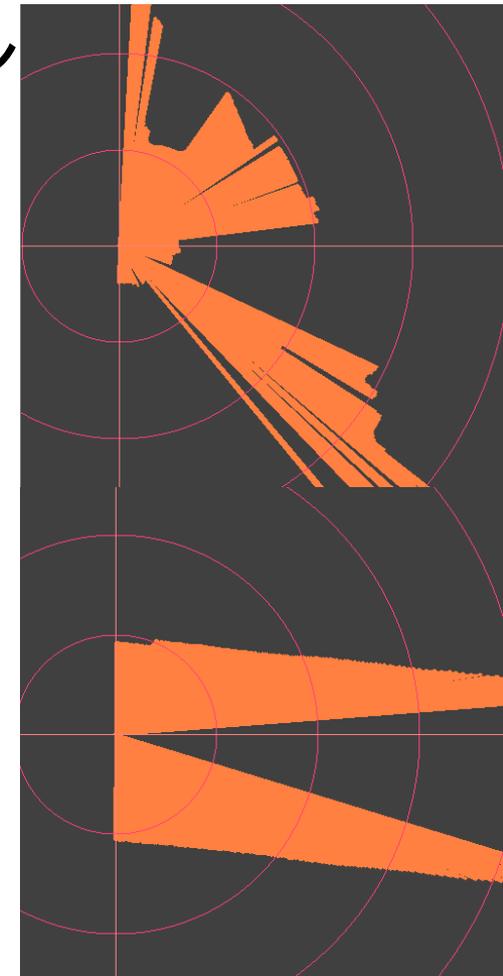
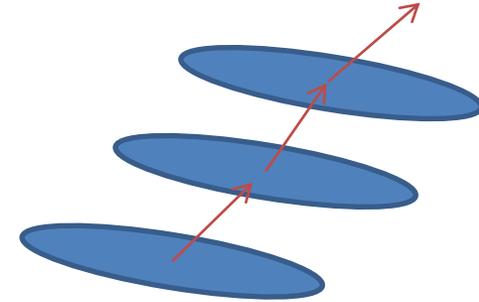
階層システム

- 下位階層で観測点をガウスクーネルでクラスタリング
- 上位階層でガウスクーネルの移動ベクトルをクラスタリング

実験:

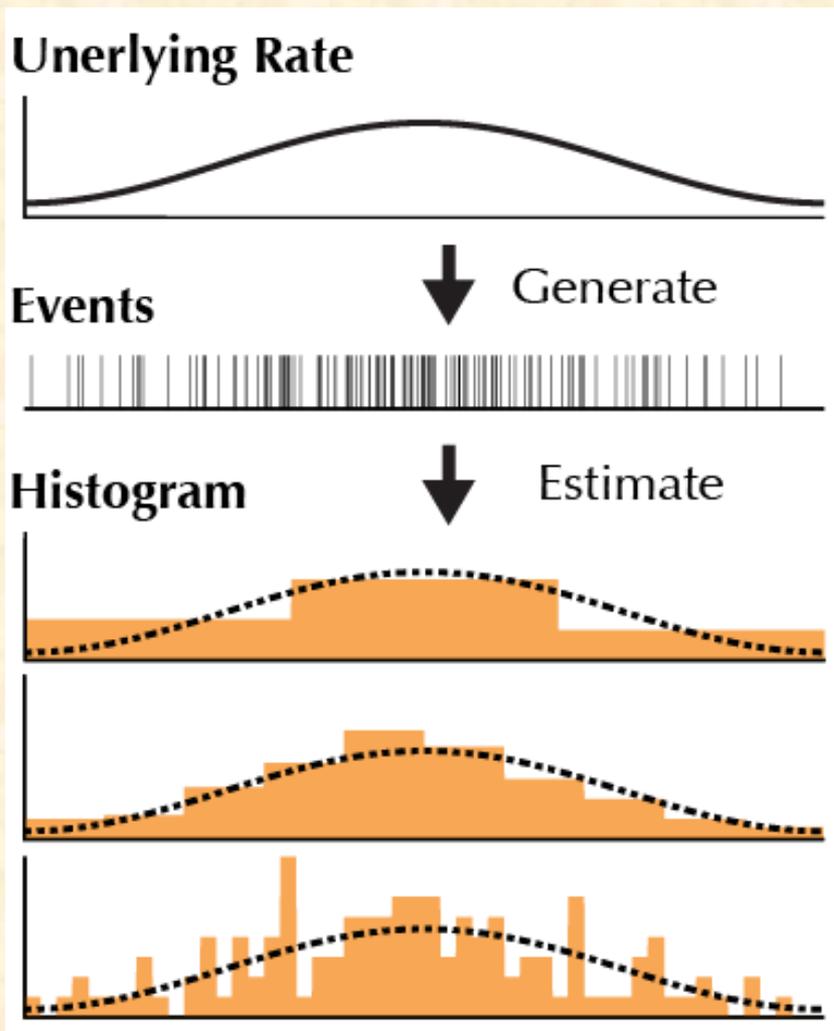
メカナムホイールロボットに3次元計測センサを搭載

上位クラスタの特性により、補正・予測・移動物体の検知など



D-123 相関のあるイベント時系列に対するヒストグラムの最適化

近江 崇宏, 篠本 滋(京大・理)



ヒストグラムのビン幅をどのように選ぶか?

先行研究

ヒストグラムと未知の発生率との平均二乗誤差を最小化する方法を考案。ただし**イベント間の独立性を仮定**。

(Rudemo 1982, Shimazaki & Shinomoto 2007)

問題点

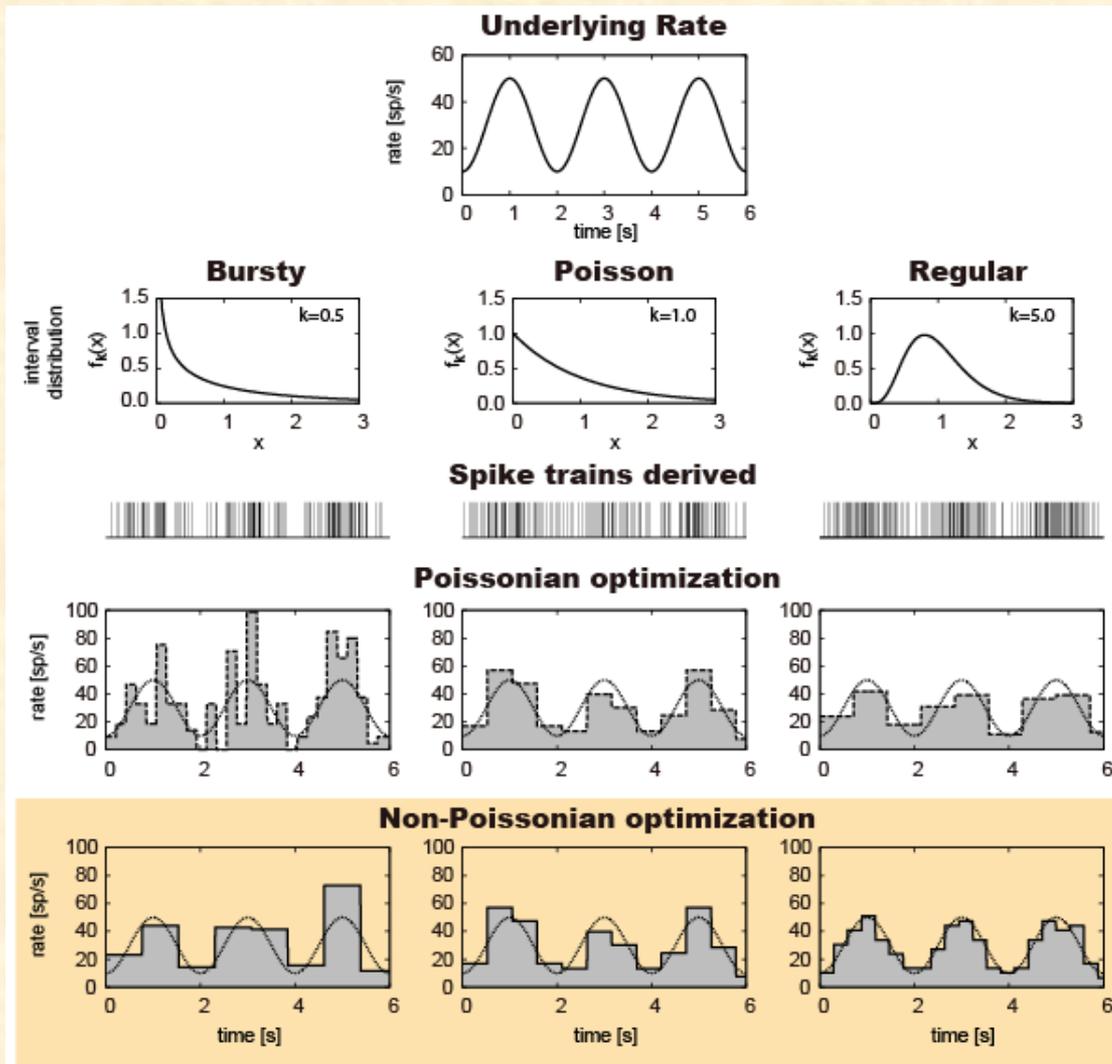
神経細胞のスパイク発火や地震などのイベント時系列には**イベント間に相関が存在**。

本研究

イベント間の相関を考慮した最適化法を提案。

(Omi & Shinomoto 2011)

数値実験を用いて推定精度が改善されたことを確認した。



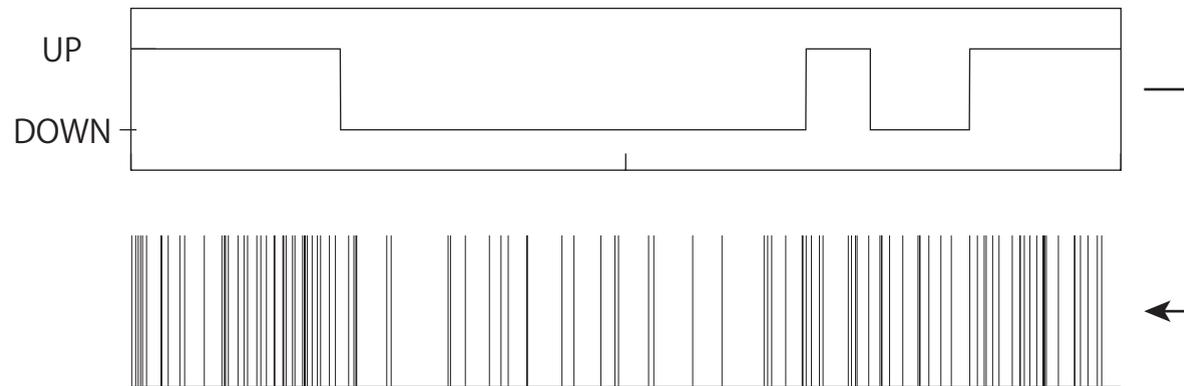
[Ref] T. Omi & S. Shinomoto, “Optimizing time histograms for non-Poissonian spike trains”,
Neural Computation (In press)

D-124 非一様ポアソン点過程のレート変動検出限界

新谷 俊了, 篠本 滋

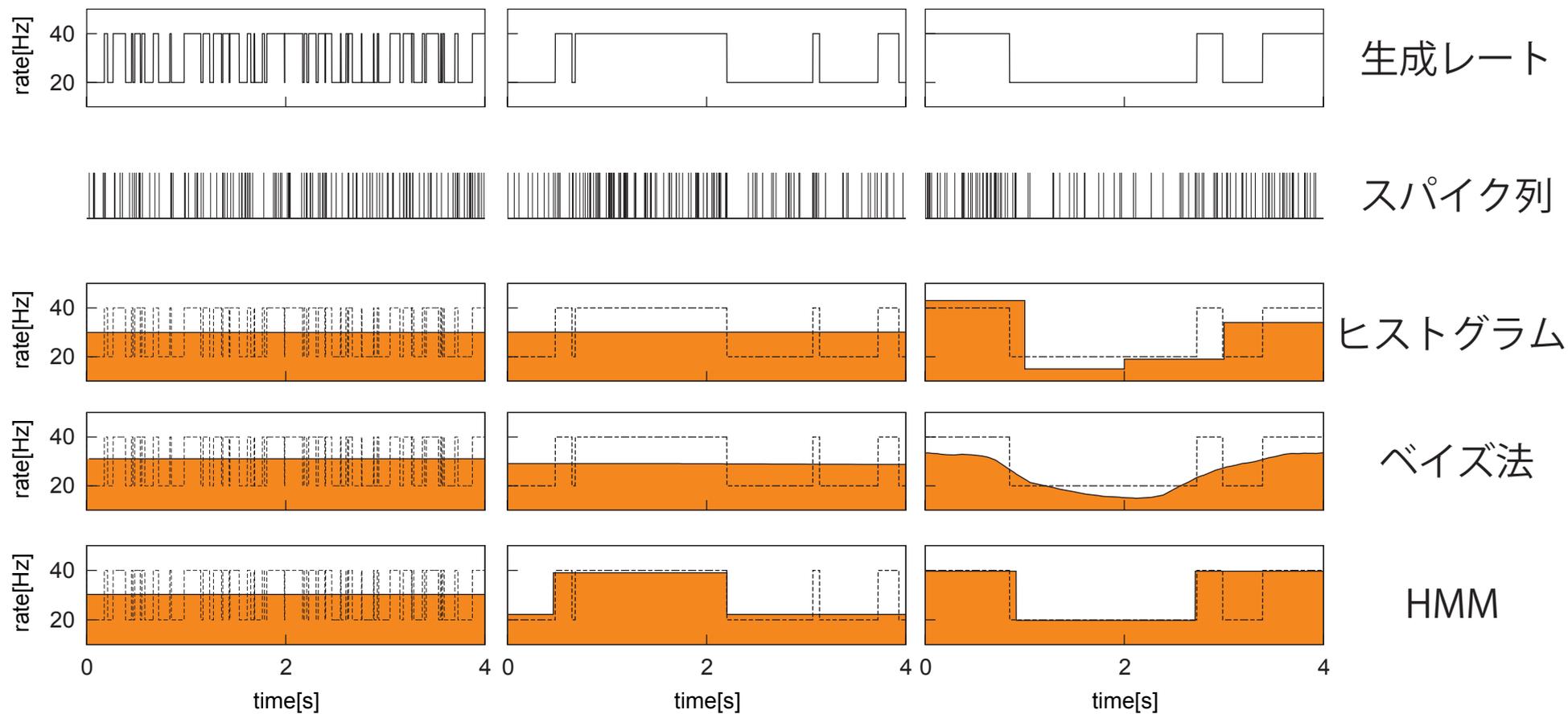
京都大学 京都大学

shintani@ton.scphys.kyoto-u.ac.jp



UP-DOWN レートから生成されるスパイク列の解析

- ヒストグラム法
- ベイズ法
- 変分ベイズ HMM



三種すべての方法でレート変動の検出限界が存在した。

D-125 改良Expert Advice を用いた時系列予測

森野佳生, 平田 祥人, 富岡 亮太, 鹿島 久嗣, 山西 健司, 合原 一幸 (東京大学)
morino_at_sat.t.u-tokyo.ac.jp

研究の目的:

複数の experts が与えられた場合に
カオス的な時系列の将来の予測を行なうこと.

提案手法:

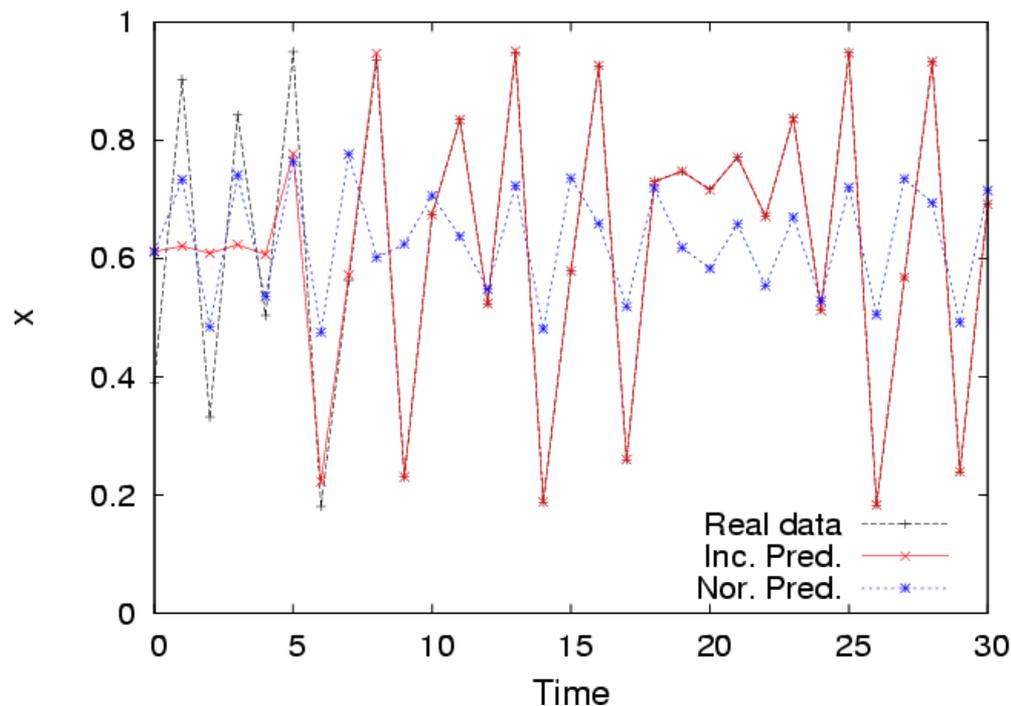
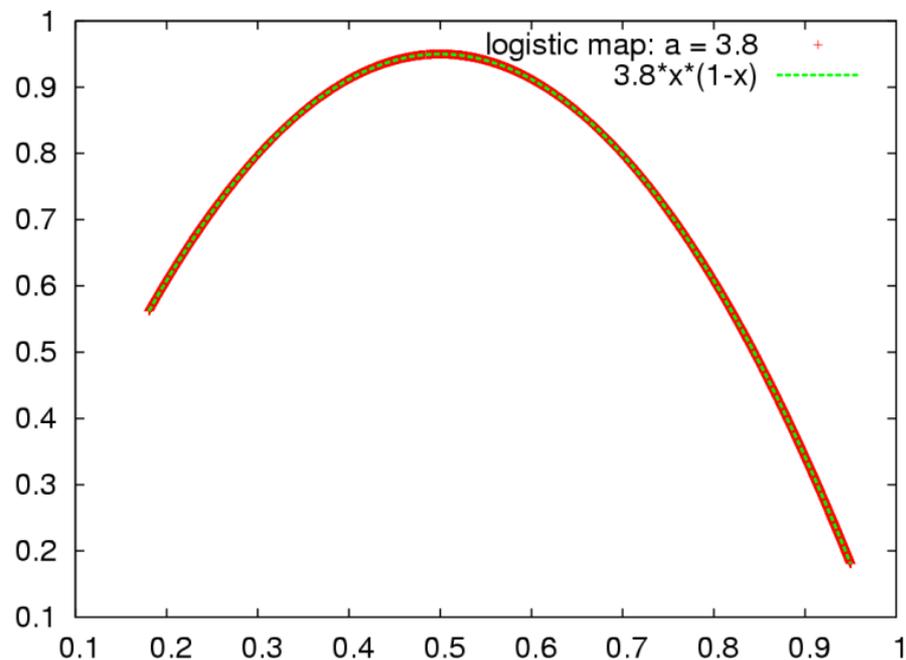
Expert advice 法に対して, 近い過去の影響を
遠い過去の影響よりも重視するように
アルゴリズムを変更する.



数理モデルと実データで有用性を確認.

数値実験結果(例):

$$\text{Logistic map: } x_{n+1} = ax_n(1 - x_n)$$



黒点 は実データ,
赤点 は今回の手法による予測,
青点 は従来手法による予測.

D-126 Compressed Sensing を利用した欠損画素補完による 仮想視点画像の画質改善

東京理科大学 保坂忠明 浜本隆之

実画像



任意視点画像
(欠損画素を
黒で表示)



任意視点画像
(補完後)



補完方法



拡大図

1. 確定画素(非欠損画素)が類似している小領域を実画像から探索する
2. その類似領域を欠損画素に埋め込む

奥行き情報とエピポーラ幾何を考慮し Compressed Sensing により実現

D-127 トップ k リストのオンライン予測

安武翔太・畑埜 晃平・瀧本英二・竹田正幸 (九州大学)

各時刻 $t=1,2,\dots,T$ のプロトコル

1 $\sigma_t \in S_n$ を予測

環境



3 損失 $d(\sigma_t, \tau_t)$ をこうむる

2 トップ k リスト $\tau_t \in S_{n,k}$ をもらう

$S_{n,k}$

n 要素から k 要素を選び, それを並べてできる順列の集合. $S_{n,n} = S_n$ と表記.

$$\text{目標: } \sum_{t=1}^T d(\sigma_t, \tau_t) \approx \min_{\sigma \in S_n} \sum_{t=1}^T d(\sigma, \tau_t)$$

アルゴリズムの
累積損失

オフライン最適解の
累積損失

kendall tau 距離

$$\sum_{1 \leq i < j \leq n} I[((\sigma(i) - \sigma(j)) \cdot (\tau(i) - \tau(j))) < 0]$$

ただし, $I[\text{真}] = 1, I[\text{偽}] = 0$

提案手法 (PermRank)

主なアイデア

- 順列を **比較ベクトル** $q = \{0, 1\}^N$ で表現
- 比較ベクトルの予測を **確信度ベクトル** $p = [0, 1]^N$ の予測へと“緩和”
→ 従来の典型的なオンラインアルゴリズムを適用可能

$$N = n(n-1)/2$$

$$\sigma = (1, 3, 2, 4)$$
$$q = (1, 1, 1, 0, 1, 1)$$

提案手法の性能

累積損失(期待値)の上界:

$$E \left[\sum_{t=1}^T d(\sigma_t, \tau_t) \right] \leq \frac{3}{2} \min_{\sigma \in S_n} \sum_{t=1}^T d(\sigma, \tau_t) + O(n\sqrt{nkT})$$

$k = n$ における, 下界:

$$E \left[\sum_{t=1}^T d(\sigma_t, \tau_t) \right] \geq \min_{\sigma \in S_n} \sum_{t=1}^T d(\sigma, \tau_t) + \Omega(n^2\sqrt{T})$$

($1+\varepsilon$)倍近似を達成する($1/\varepsilon$ に関する)多項式アルゴリズムは($NP \not\subseteq BPP$ の仮定のもと)存在しない。よって, 提案手法はほぼ最適。

実験結果(実データ)

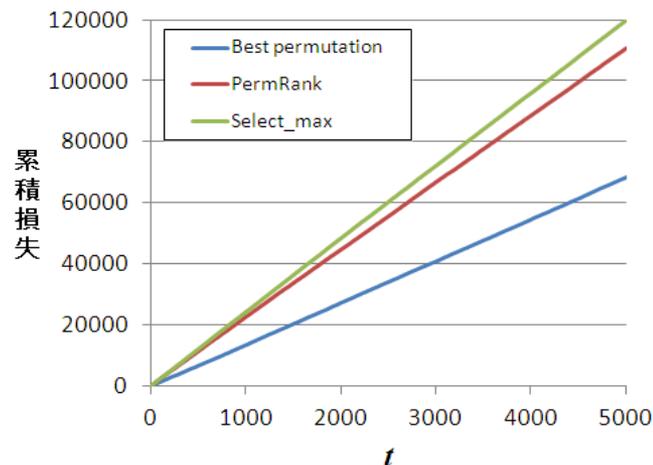
$n=10, k=7$ で実験を行った。

素朴な手法(Select_max)よりもよい結果が得られていることがわかる。

Select_max ...

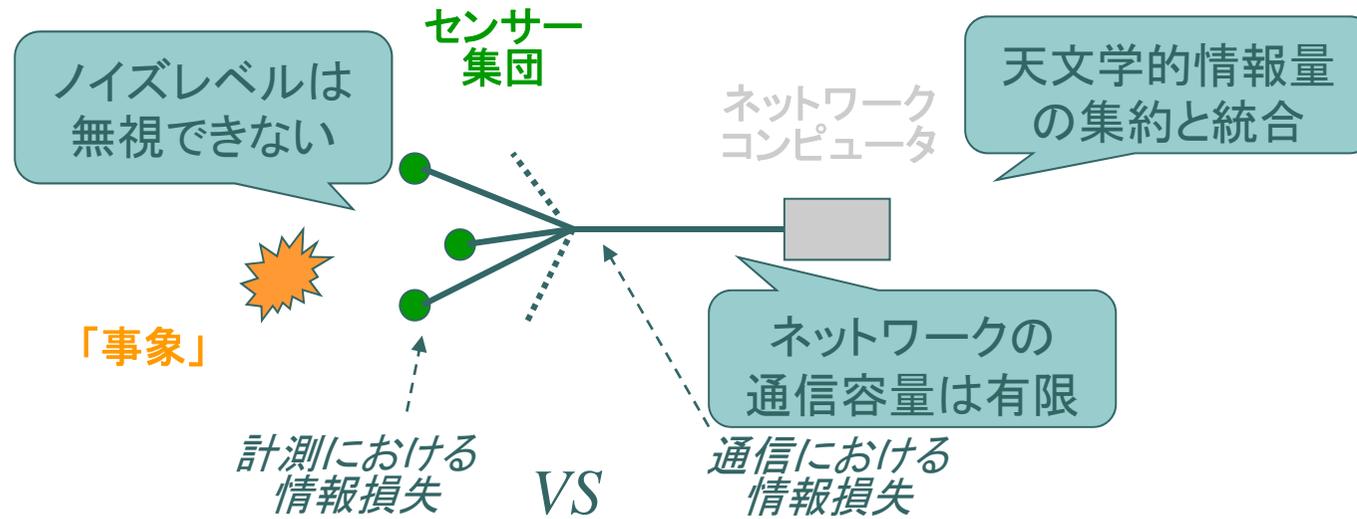
過去における出現頻度が高い k 要素を出力

累積損失の比較



《センサーネットワーク》

ひとつの「事象」を多数のセンサーを使って計測し、そのデータを集めることで、できるだけ正確にそこで何が起きているのかを推定したい。



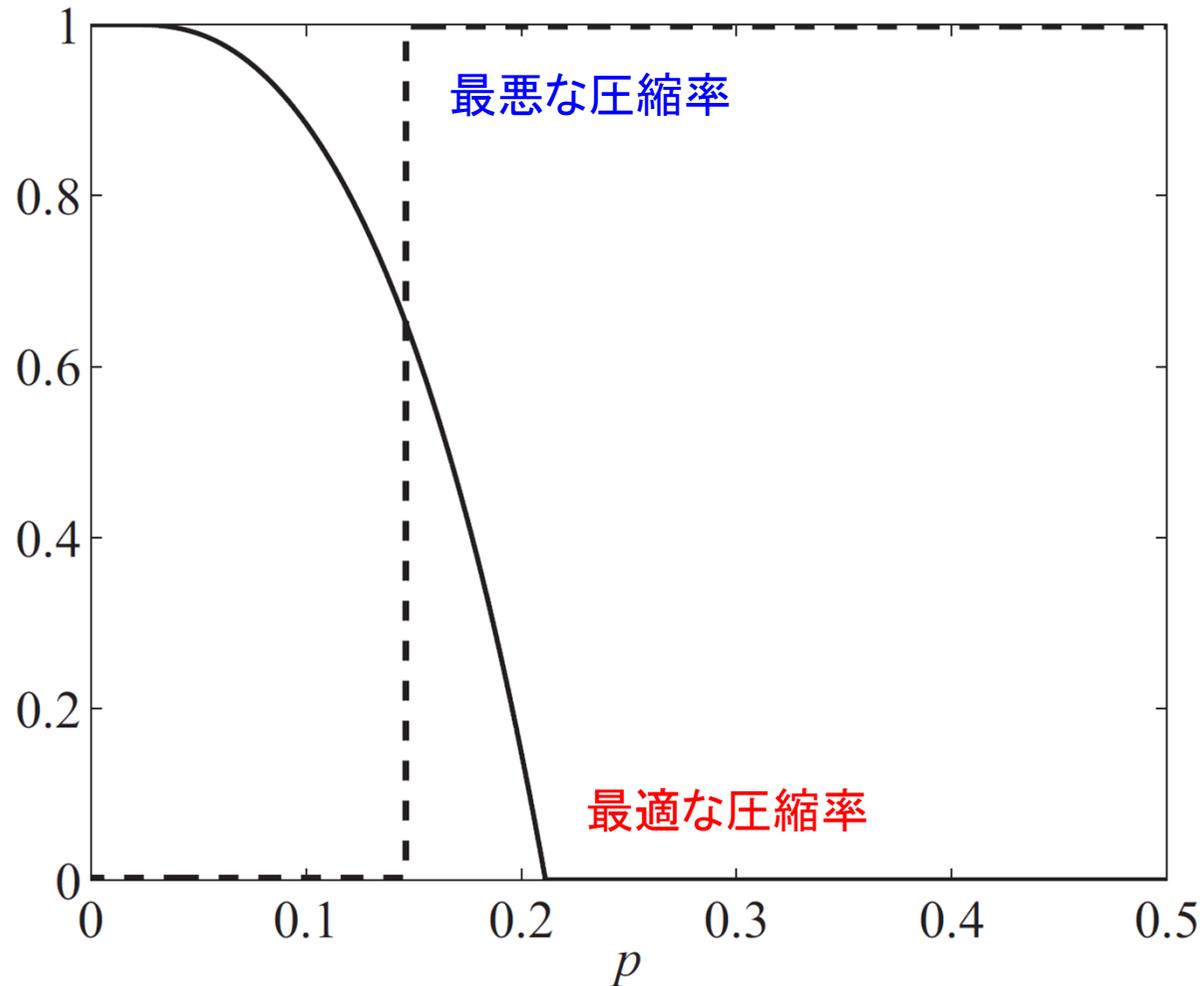
高ノイズ領域

ネットワークが広域化・高密度化することで適用範囲が拡大！

有限容量

センサー情報をうまく圧縮しないと帯域を有効利用できない！

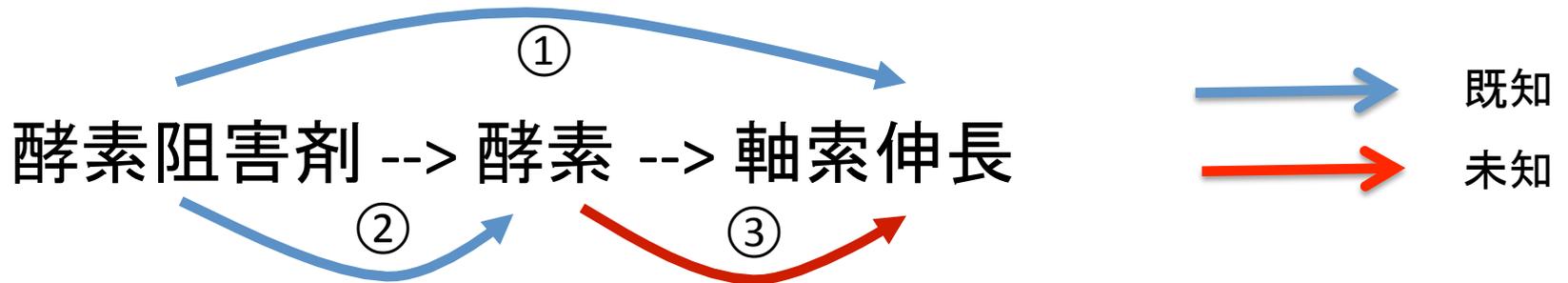
《課題》 ネットワーク容量が有限のとき、最適な圧縮率は定義できるか？



- ・ネットワーク容量が**無限に発散**する極限を議論する。(当然, すべて見えてしまう.)
- ・圧縮率のノイズレベルに対する最適性は, データ統合による誤り率が**ゼロに収束していく速度の大きさ**で測ることにする.

D-129: 圧縮センシングによる神経軸索の伸長を制御する因子の同定に関する検討

丸野由希, 宮本敦史, Hassan Al-Ali, Vance Lemmon, 作村諭一, 池田和司



① 軸索伸長度

阻害剤A	2
阻害剤B	-1
⋮	0
	5

② 酵素1 酵素2 ...

阻害剤A		阻害			阻害	
阻害剤B	阻害		阻害		阻害	
⋮						阻害
				阻害		

③

酵素1	?
酵素2	?
酵素3	?
⋮	?

目的: 効率良い軸索伸長をする”酵素阻害の組合せ”を同定

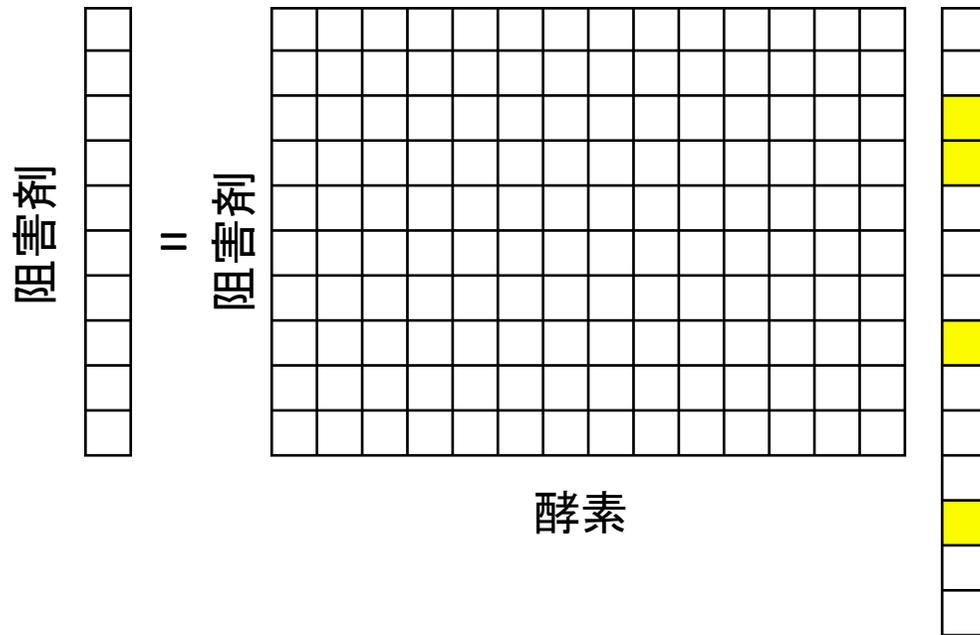
圧縮センシングによる定式化

仮定：軸索伸長を邪魔する酵素はスパースである

y : 観測伸長度

A : 伸長貢献度

x : 重み



$$y \in R^m$$

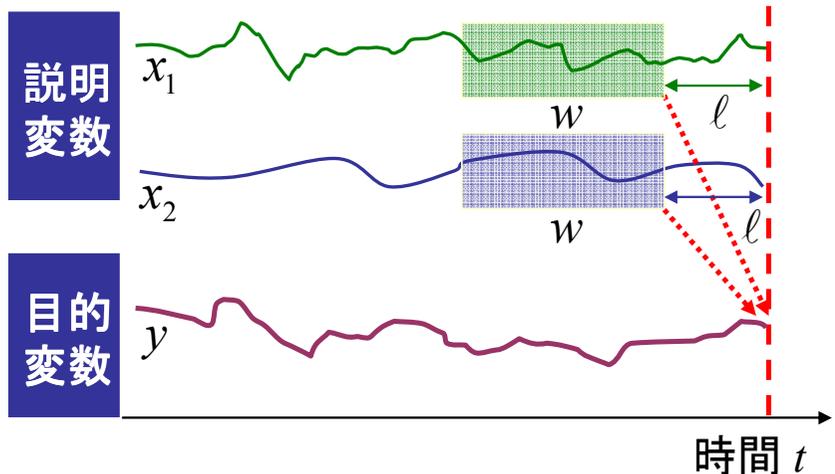
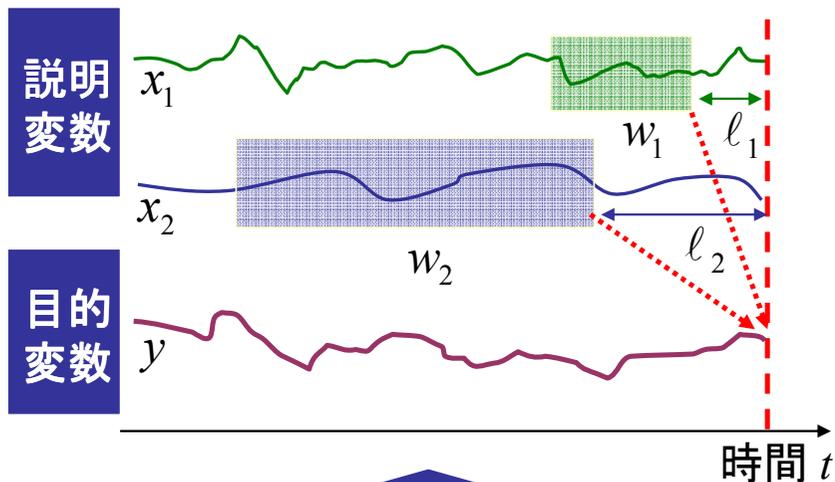
$$A \in R^{m \times n} (m < n)$$

$$x \in R^n \quad \|x\|_0 = k$$

$$\min \|x\|_1 \quad \text{s.t.} \quad Ax = y$$

重み x を求める事により,
軸索伸長に寄与する”酵素阻害の組み合わせ”を同定する

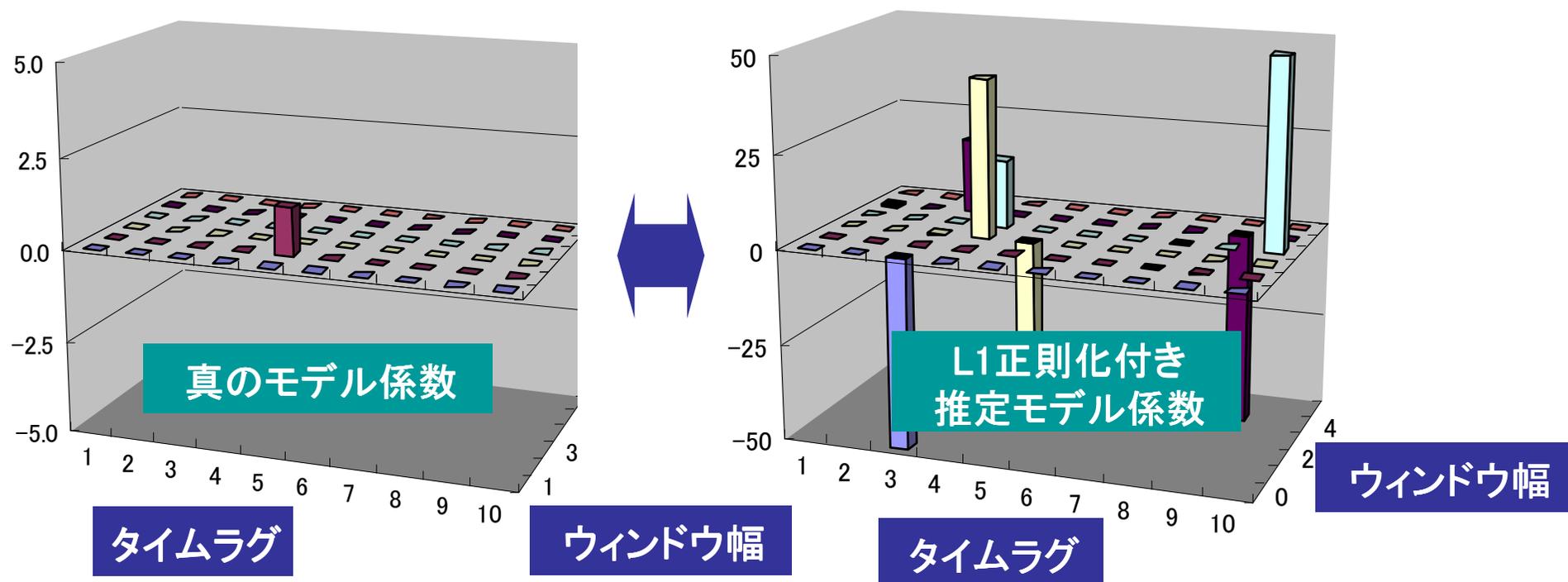
[D-130] 多様なタイムラグを含む 時系列予測のための変数選択法 比戸 将平(IBM)



- 時系列予測の例
 - 金融データ・経済指標
 - 医療センサーデータ
 - 機械系センサーデータ
- 目的変数との間の複雑な時間因果関係
 - 多様なラグ
 - 多様なウィンドウ幅

⇔ラグやウィンドウ幅は固定するのが一般的

[D-130] 多様なタイムラグを含む 時系列予測のための変数選択法 比戸 将平(IBM)



網羅的なラグxウィンドウの組では過学習 & 高コスト
⇒ラグとウィンドウを一度に選択する効率的な手法は？

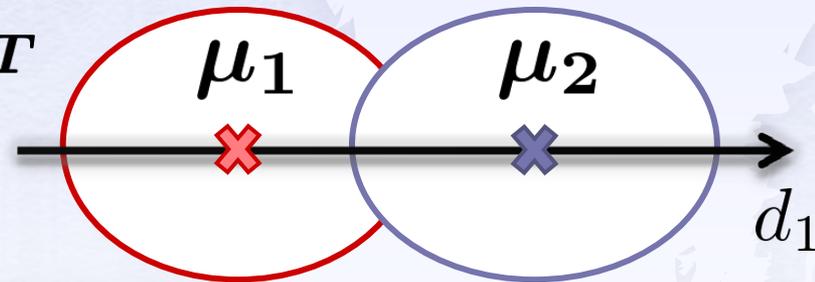
D-131 仮想クラス生成に基づく 正準相関分析への特徴追加

大橋 司[†], 木村 昭悟^{††}, 坂野 鋭^{††}, 澤田 宏^{††}

[†]同志社大学, ^{††}NTT

$$S_B x = \lambda S_W x$$

$$S_B = \sum_{i=1}^C (\mu_i - \mu)(\mu_i - \mu)^T$$



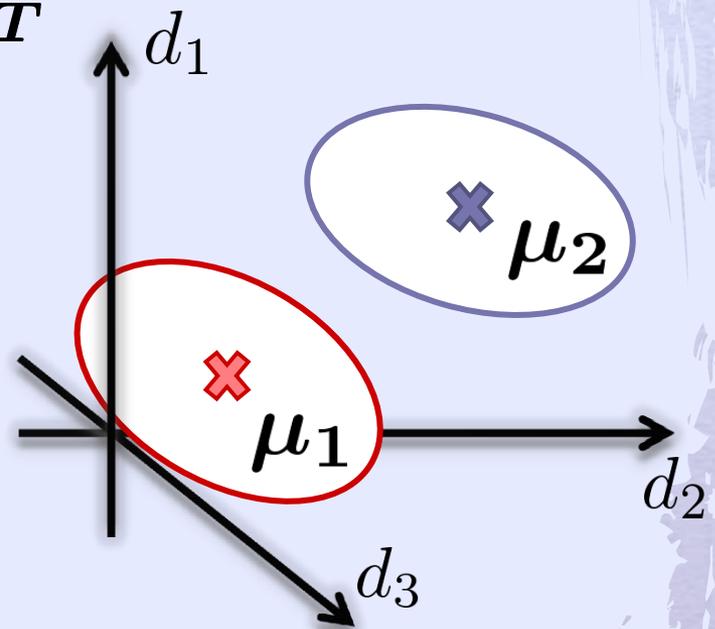
判別空間の次元数がクラス数未満に

従来法: 直交判別分析、判別空間の補空間の利用など

「仮想クラス」 新しい形の特徴追加法

$$C + C(\text{virtual})$$
$$S_B = \sum_{i=1} (\mu - \mu_i)(\mu - \mu_i)^T$$

仮想クラス分 $C(\text{virtual})$ を追加して
クラス数そのものを増加



どのような仮想クラスを作るのか？
このアプローチで識別性能を上げるには？

D-132

Just-In-Timeモデリングを用いた 太陽光発電将来出力予測

NTTコミュニケーション科学基礎研究所

藤原 幸一・須山敬之



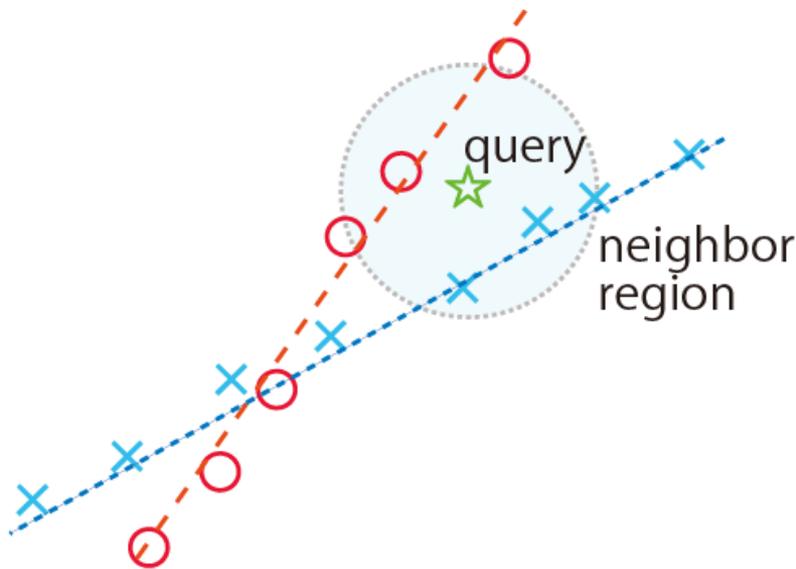
再生可能エネルギーにおける様々な変動

- 短期的：天候の急激な変化
- 中期的：季節変動
- 長期的：装置の性能劣化



様々な変動に対応した太陽光発電
将来出力予測モデルの開発

Just-In-Timeモデリングの利用



4時間先まで予測可能

D-133 圧縮センシングに基づく非線形モデリング及び弱制限等長性との関連

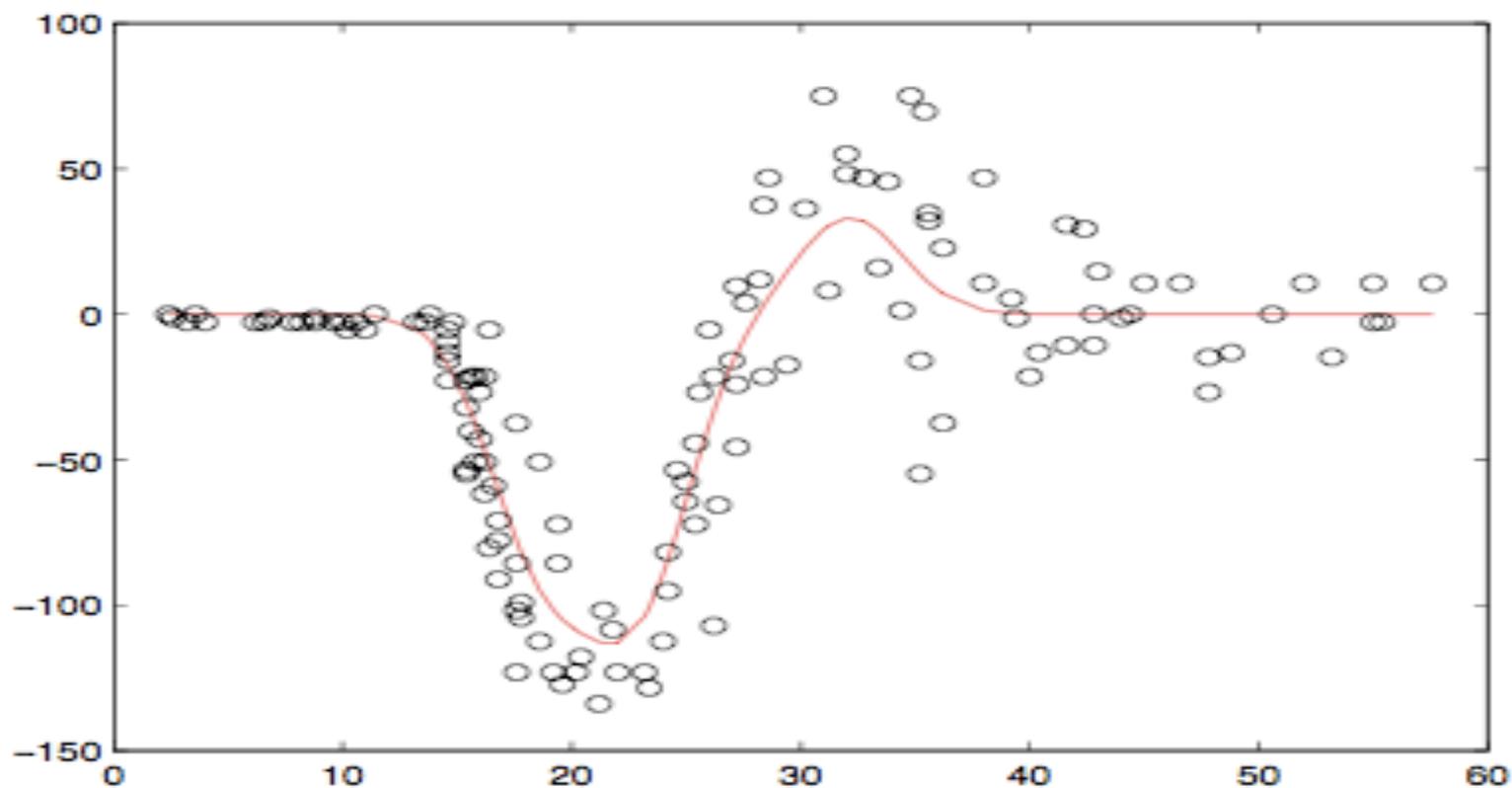
九州大学大学院数理学府 井上 寛

h-inoue@math.kyushu-u.ac.jp

本報告は圧縮センシングの考え方を基に回帰問題に適用する。特に、曲線フィッティングに着目し、不良設定問題から生じる推定の不安定性に対し、推定精度の向上を試みる。また、提案手法の数学的背景において非常に重要である制限等長性 (restricted isometry property; RIP) の条件を緩めた、弱制限等長性 (weak RIP) という概念を明確に定義し、新しく得られた結果を紹介する。

実データ（オートバイ衝突実験データ）

オートバイ衝突実験データに提案手法を適用する.このデータは,133台のオートバイを試験的に衝突させ,衝突からの経過時間 X (ms)における人体頭部への加速度 Y (g)を計測したものである.



[D-134] 確率的生成モデルによる 位相保存写像のアルゴリズム導出

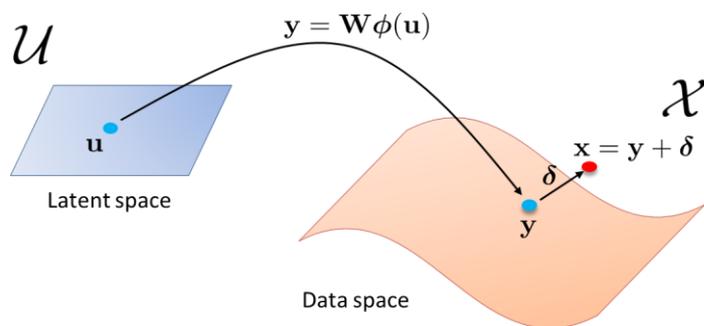
松下 聡史 中野 将秀 古川 徹生

九州工業大学大学院 生命体工学研究科 脳情報専攻

本研究の目的

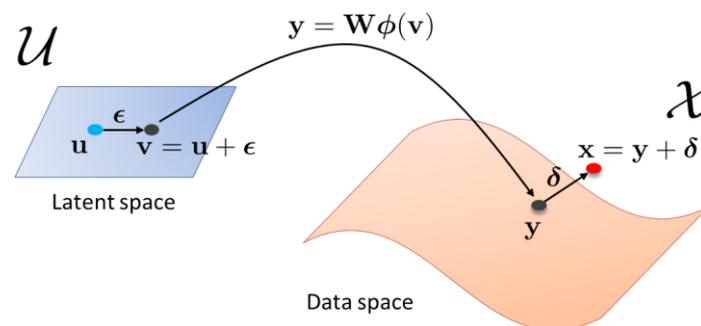
Self-organizing Maps (SOM) や Generative Topographic Mapping (GTM) を
始めとする位相保存写像群を包括するカノニカルな理論の確立

従来の生成モデル



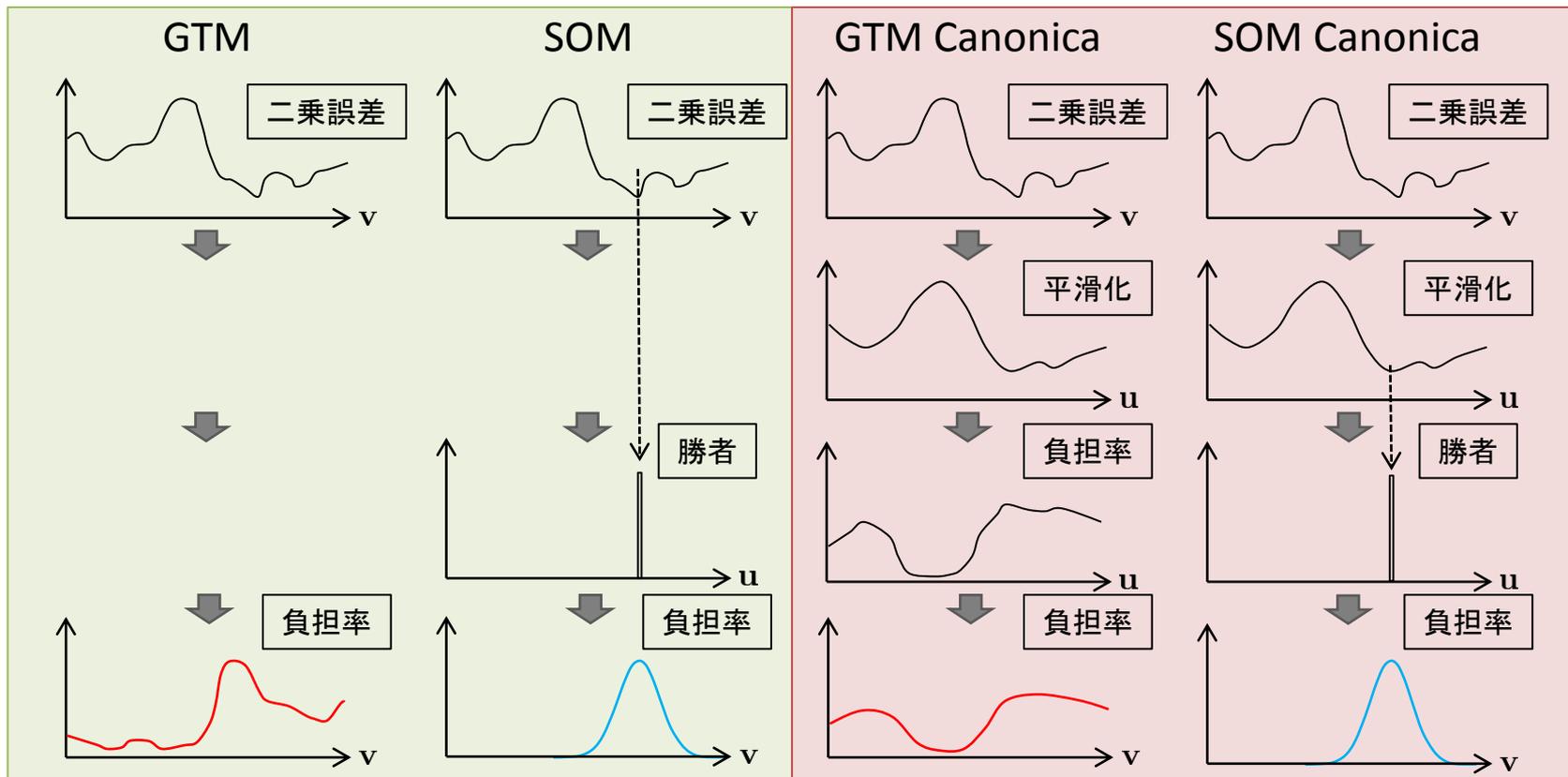
位相保存は写像を拘束することによって達成される

本研究で用いる生成モデル (カノニカルモデル)



潜在ノイズが潜在空間の位相を表現する

各アルゴリズムの E-step 比較



【特徴】

- ◆ 自然な形で位相保存の制約を与える**潜在ノイズ**の導入
- ◆ **変分ベイズ法**を用いた潜在変数, 各種パラメータの推定
- ◆ **SOM と GTM の統一的理解**
- ◆ グラフ構造など離散表現への柔軟な拡張 (NG Canonica)

D-135 強度と位相を変化させる結合によるネットワークの提案

城 真範, 赤穂 昭太郎

(産業技術総合研究所 ヒューマンライフ 情報数理)
shiro@ni.aist.go.jp

- 従来のニューラルネットは、重み付き入力の加算にSigmoid関数をかませて非線形表現を実現→カオス時系列データでうまくいかない例があった。

一般の時系列では、遅れ時間の取り方とモデルの次元がどちらも未知である。
(埋め込み次元の推定法は存在するが、モデルの次元は分からない)
→遅れ時間もモデルの次元に組み入れて、超高次元モデルにしてしまい、lassoを使って不必要な係数を除去することで、素早く最適モデルを作る。

- 定式化

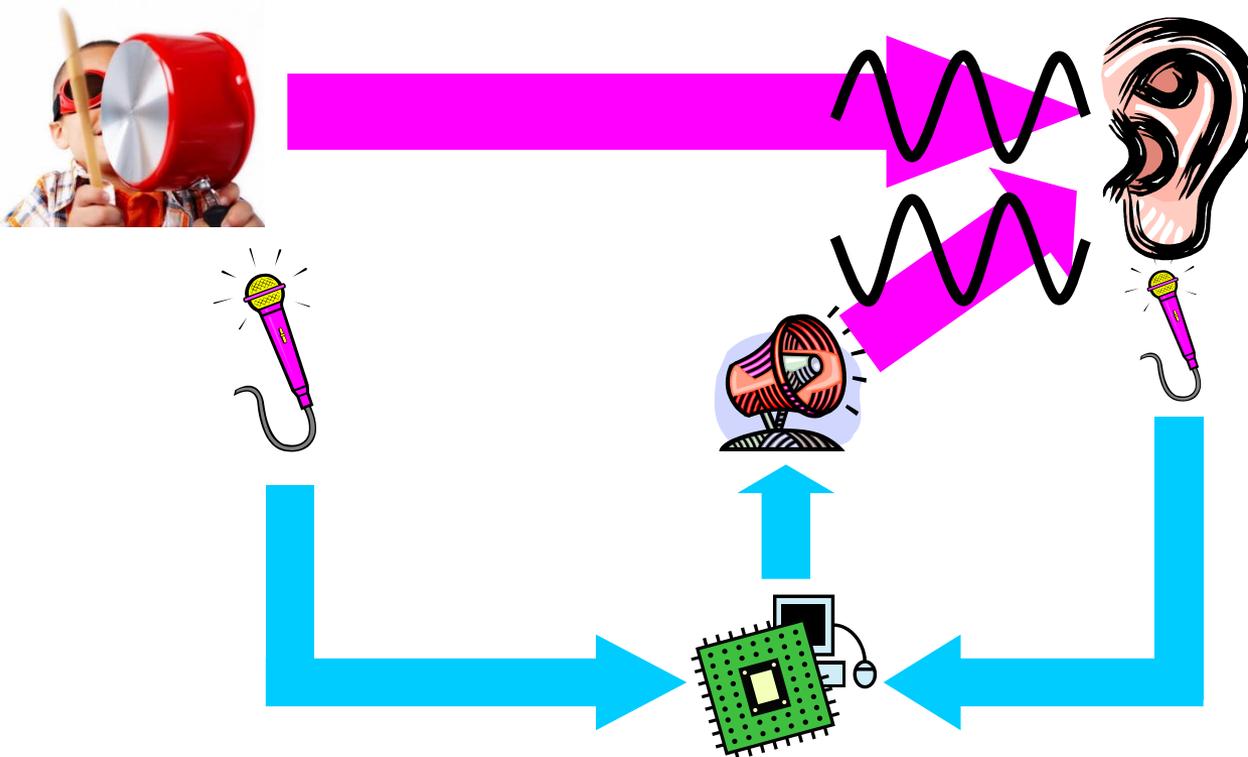
$$E_t \left(x_{t+1} - \sum_{i=0}^{i_{\max}} \sum_{j=0}^{j_{\max}} a_{ij} x_{t-i}^j \right)^2 + \lambda \sum_{i=0}^{i_{\max}} \sum_{j=0}^{j_{\max}} w_{ij} |a_{ij}|$$

- 結果:ポスターでご覧ください。

統計力学的手法による適応信号処理の解析

三好 誠司(関西大) 梶川 嘉延(関西大)

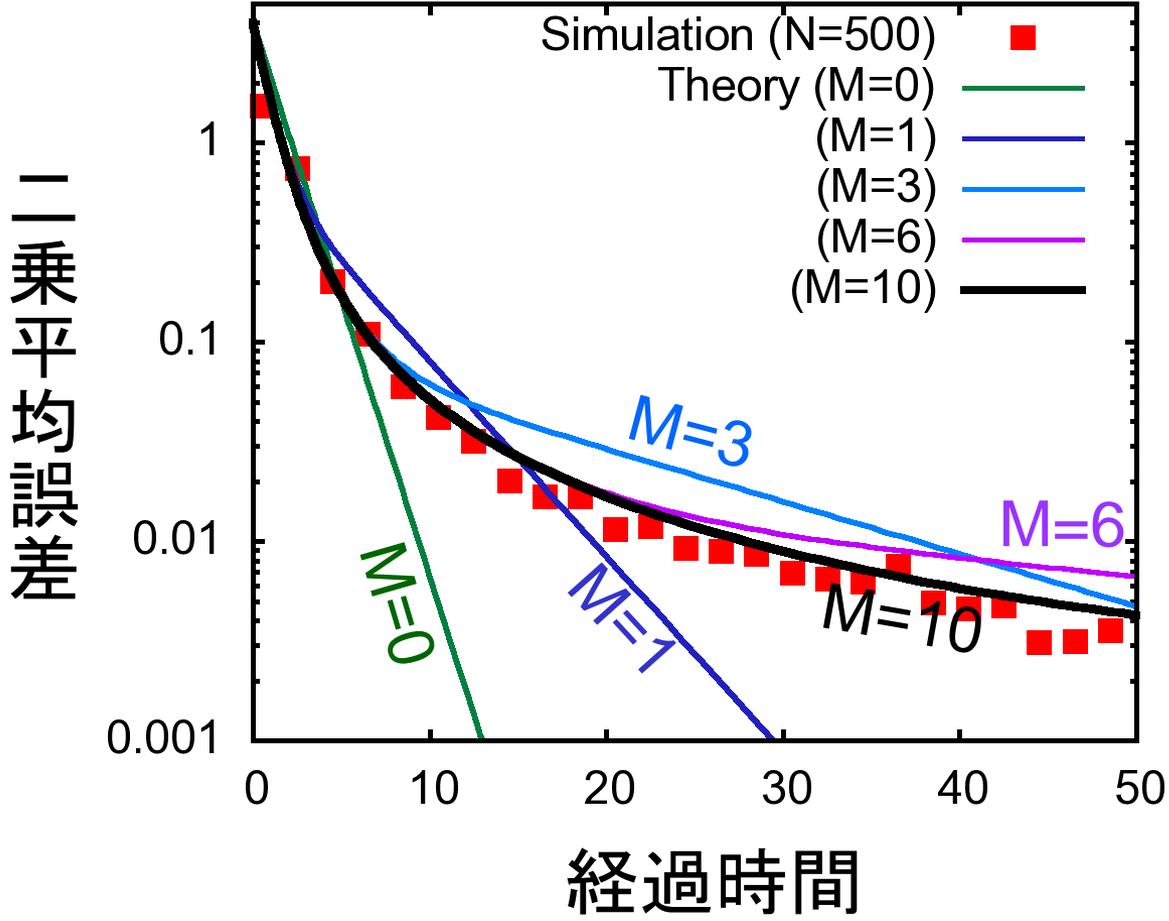
雑音源



広く使われているにも関わらず
一般的な理論は無い！

統計力学的手法による適応信号処理の解析

三好 誠司(関西大) 梶川 嘉延(関西大)



D-137 不完全 Gibbs sampling(高畠 一哉 産総研ヒューマンライフテクノロジー)

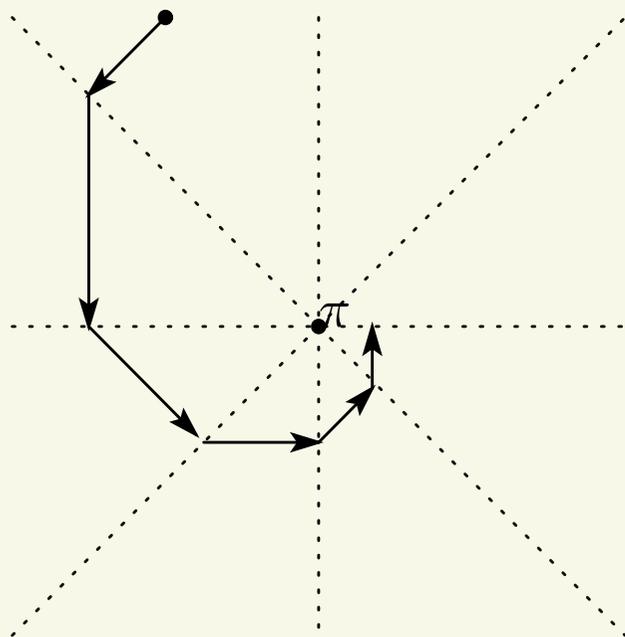
$X = (X_0, \dots, X_{n-1})$: 複数の確率変数

$\pi(X)$: 目標となる分布

$\pi(X_i|X_{-i})$: 全条件分布. X_{-i} は X_i を除く全ての確率変数.

Gibbs sampling では逐次巡回的に (あるいはランダムに) i を選び, 条件付分布 $\pi(X_i|x_{-i})$ から新たな x_i をサンプリングすること (ノード i の発火) を繰り返す. 出てくるデータの分布は $\pi(X)$ に収束する.

もし X_{-i} 中の全変数を見るのを止めてその一部 Y_i しか見ない, つまり $\pi(X_i|x_{-i})$ からではなく $\pi(X_i|y_i)$ からサンプリングを行うとデータの分布はどこに収束するだろう?



完全 Gibbs sampling

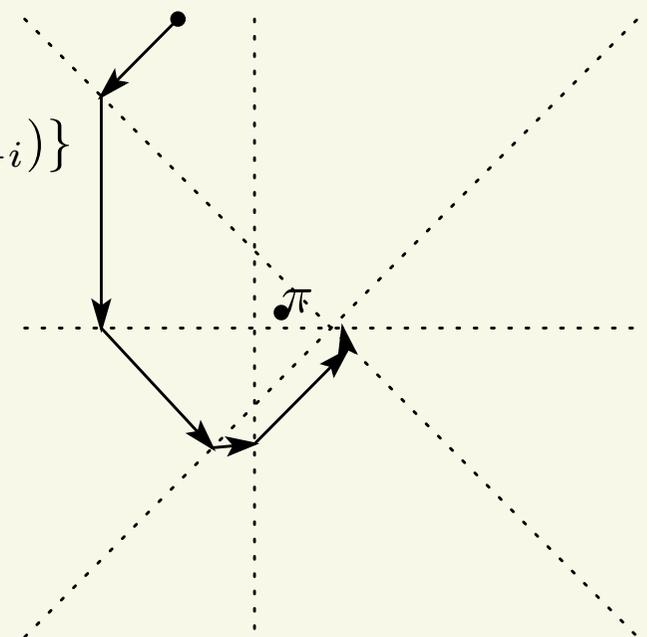
左図点線は E^b 多様体

$$\{p(X) | p(X_i|X_{-i}) = \pi(X_i|X_{-i})\}$$

右図点線は E^b 多様体

$$\{p(X) | p(X_i|Y_i) = \pi(X_i|Y_i)\}$$

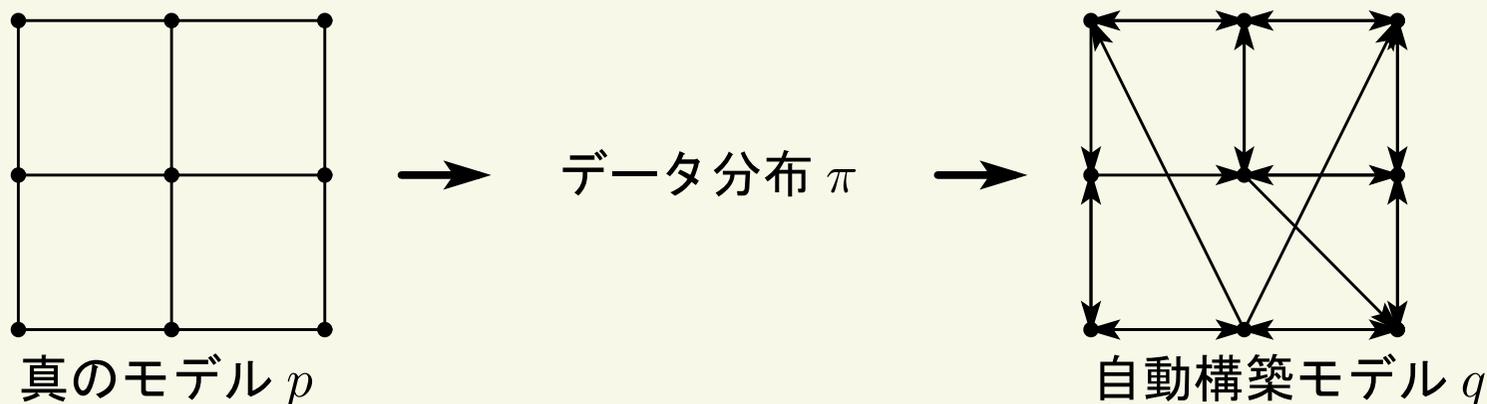
矢印は m 射影



不完全 Gibbs sampling

D-137 不完全 Gibbs sampling

不完全 Gibbs sampling を使うとデータから確率モデルを自動的に構築する仕組みが作れる.



9スピンのイジングモデル $p(X)$ (X は 9 ビットの 2 値データ, $|X| = 512$) から 100 個の iid データをサンプルし (データ分布 $\pi(X)$) それを不完全 Gibbs sampling による自動学習器に見せてモデル $q(X)$ を自動構築させたところ

$$KL(\pi||p) = 0.72 \quad KL(\pi||q) = 0.49 \quad KL(q||p) = 0.37$$

$$KL(\pi||q), KL(q||p) < KL(\pi||p) \quad !!$$

D-138 ベイズ推定に基づくX線CT再構成でのビームハードニング低減

牧野 貴樹*,水谷 治央*,小塚 淳**,佐藤 一誠*

*東京大学 **大阪大学

- CT:2次元断面像を1次元投影像の組から再構成

投影行列

断面像 (未知)

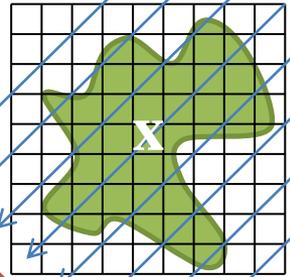
投影像 (既知)

$$y_i = W_i \mathbf{x} + \varepsilon$$

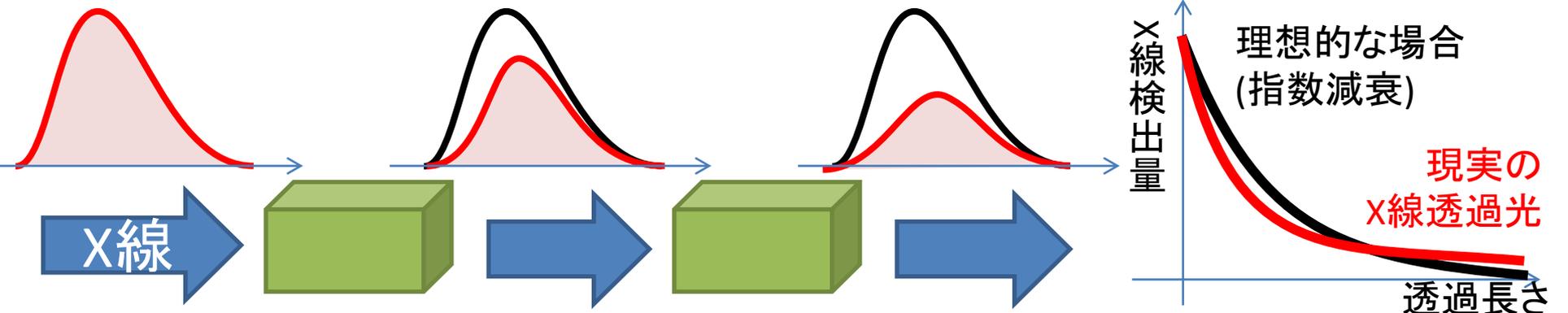
理想的な場合の式

ノイズ

y_i



- ビームハードニング: X線スペクトルが物体の透過に伴い高エネルギー側にシフト



- 物質によって吸収係数比が異なるため、単純な補正ではアーティファクト (偽像) が消えない

→ 各画素ごとの密度 $x_{i,j}$ および吸収係数の比 $z_{i,j}$ を推定する問題をMAP推定の形で定義

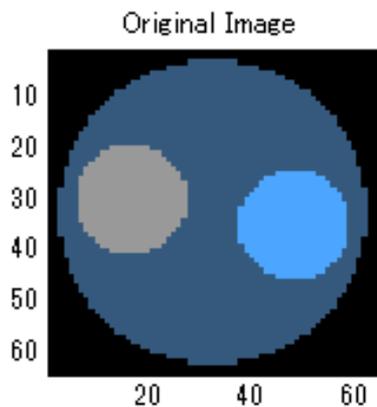
$$\mathbf{y}_i = \log\left(\sum_{c=1}^3 \lambda_c \exp W_i \mathbf{x}_c + \varepsilon\right)$$

$$x_{i,j}^{(c)} = x_{i,j} (1 + (c-2)z_{i,j})$$

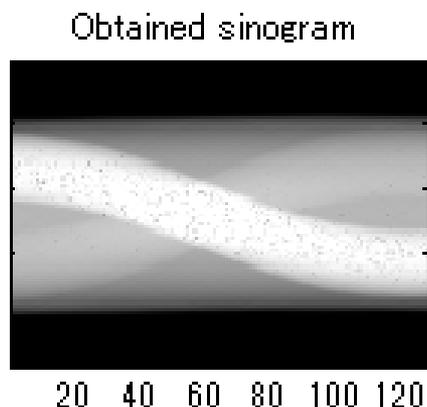
縁が明るくなる

放射状の偽像

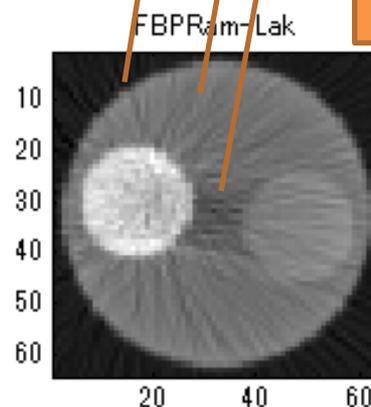
低吸収率の領域が見える



元の断面像 (色で級数係数の比を表わす)



サイノグラム (投影像)



従来法 (FBP法)



提案法

D-139 マハラノビス距離における主成分展開の研究

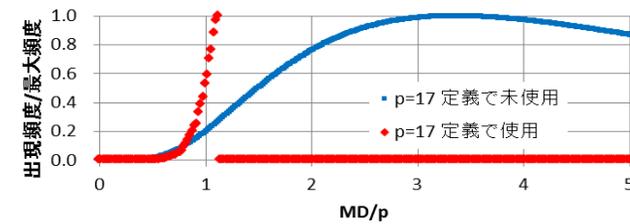
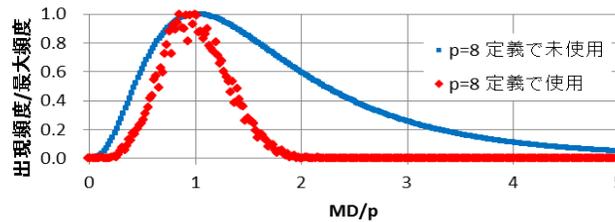
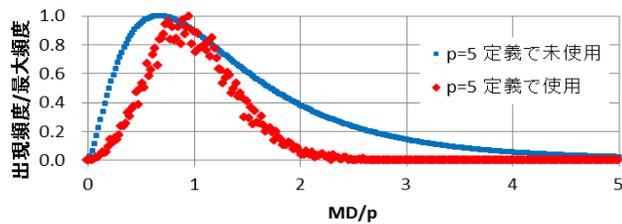
小林 靖之(帝京大学 理工学部)

ykoba@ics.teikyo-u.ac.jp

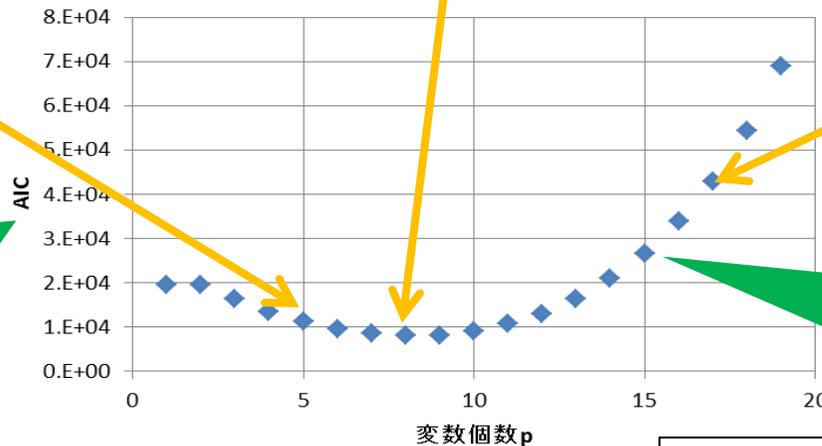
概要

- マハラノビス距離の定義に用いる相関行列の標本データ個数 n と変数個数 p に、マハラノビス距離(MD)は次の通り依存
 - ① 定義に用いた、基準の標本の MD :
 $p \rightarrow n - 1$ で $MD/p \rightarrow 1$ へ非対称分布を取りながら収束
 - ② 定義に用いなかった、基準の標本の MD :
 $p \rightarrow n - 1$ で 期待値 $E(MD/p) \rightarrow \infty$
 - ③ 定義に用いなかった、基準の標本の MD の従う分布に最も適合する変数個数 p_{AICmin} は、標本データ個数 n に依存 ($p_{AICmin} \sim \sqrt{n}$ ($n \leq 200$) に従う?)
- 従って、 MD を主成分で展開し、③の関係を満たすように主成分個数を制限すれば、定義への採用如何を問わず標本データを MD で正確に識別できる。

モンテカルロシミュレーションによるマハラビス距離(MD)の度数分布

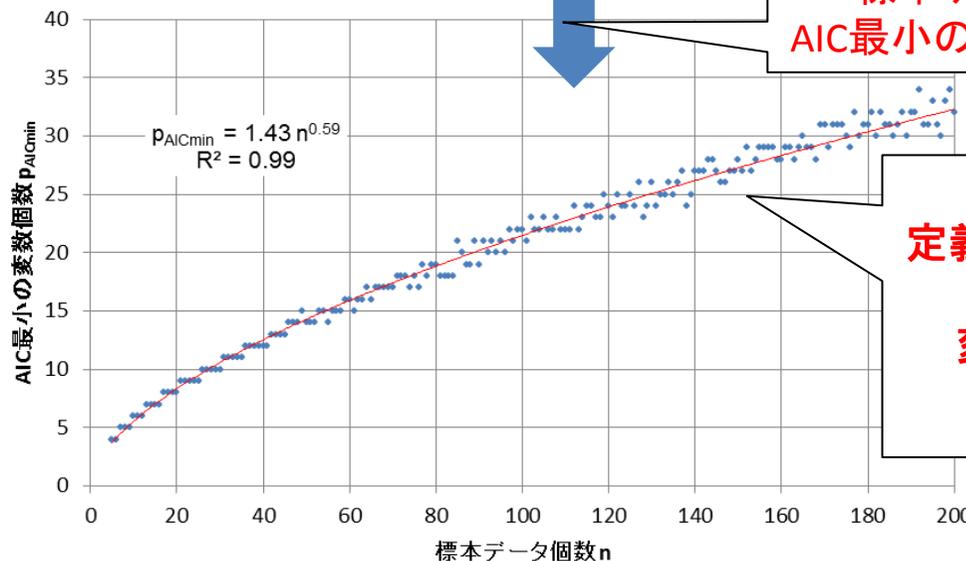


定義に不採用の
基準標本のMDが
従う分布(~F分布)を元に、
定義に採用したMD標本
に関するAICを計算



標本データ個数 $n=20$ では、
変数個数 $p=8$ で、
定義に採用と不採用の
基準標本の
MD分布が最も近づく
(AIC最小)

標本データ個数 n に対する
AIC最小の変数個数 p_{AICmin} を計算



標本データ個数 n に対して
定義に使用・未使用の基準標本の
MD分布が最も近づく
変数個数 p_{AICmin} を求めると、
 $p_{AICmin} \sim \sqrt{n} (n \leq 200) ?$

[D-140] 位相保存写像を用いた 関係データの非線形テンソル分解

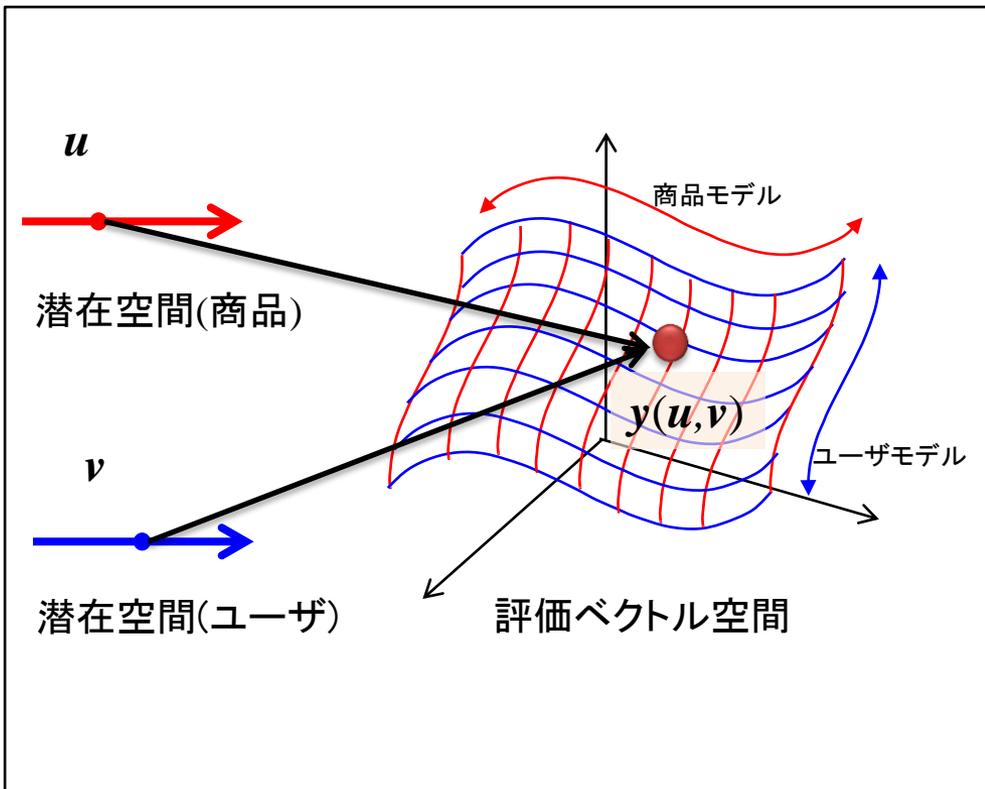
岩崎 亘, 和田沙織, 古川徹生

九州工業大学大学院 生命体工学研究科 脳情報専攻

	商品 1	...	商品 j	...	商品 J
ユーザ 1					
ユーザ 2					
⋮					
ユーザ i			x_{ij}		
⋮					
ユーザ I					

多次元評価ベクトル

$x_{ij} = (\text{評価項目1}, \text{評価項目2}, \text{評価項目3}, \dots)$



$$y(u, v) \approx W \times_1 \Phi^1(u) \times_2 \Phi^2(v)$$

↑
評価

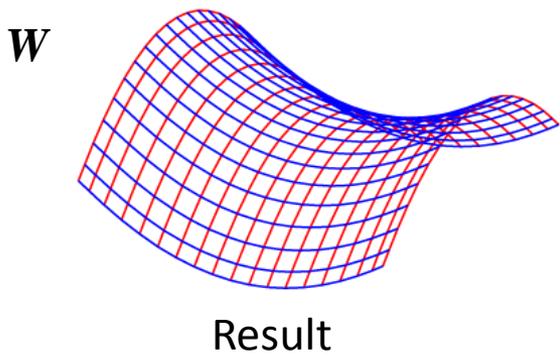
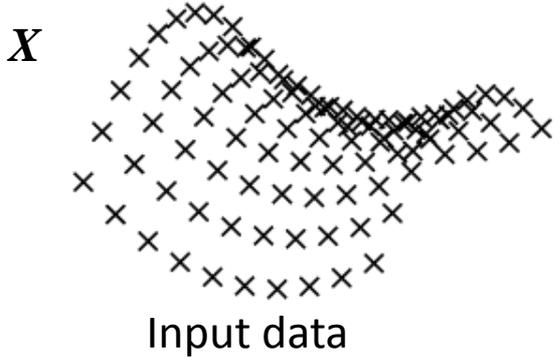
↑
ユーザ

↑
商品

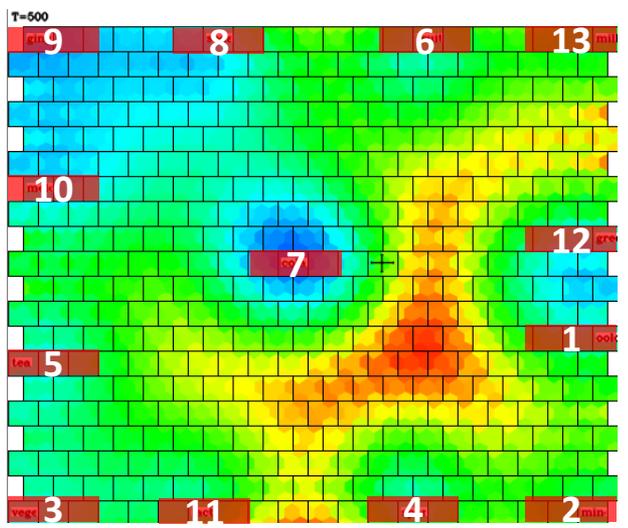
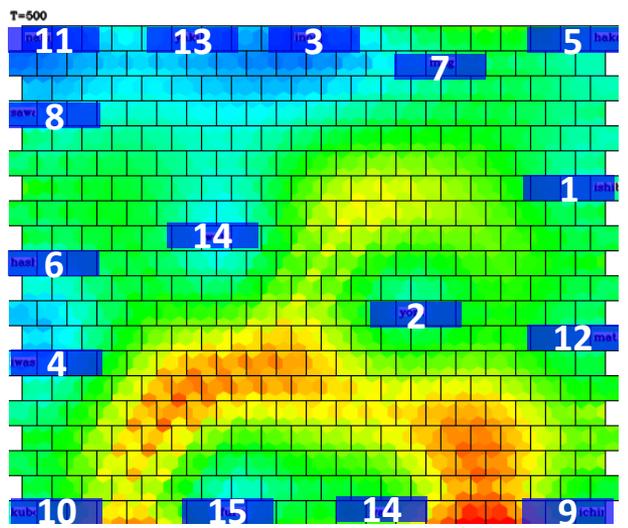
W, u_i, v_j を求める

高階自己組織化マップをもとに、非線形テンソル分解を実現

課題1:人工データ



課題2:商品評価アンケートデータ



D-141

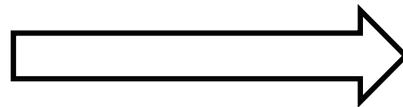
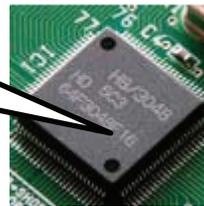
ユニット数に上限を持つカーネルマシンの重要度重みつき追記学習

山内康一郎(中部大学)

目的:リソースに制限のある組み込み機器向けカーネルマシンの開発

適用例:

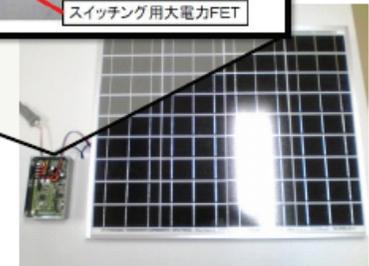
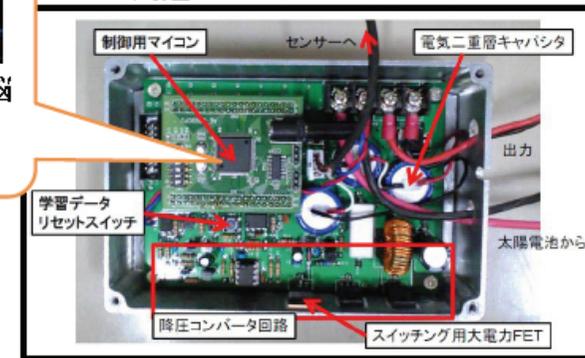
追記学習
アルゴリズム
(one-pass
Learning 可能)



マイクロコンピュータ
(※デスクトップコンピュータではない)



高速MPPTコンバータ試作機

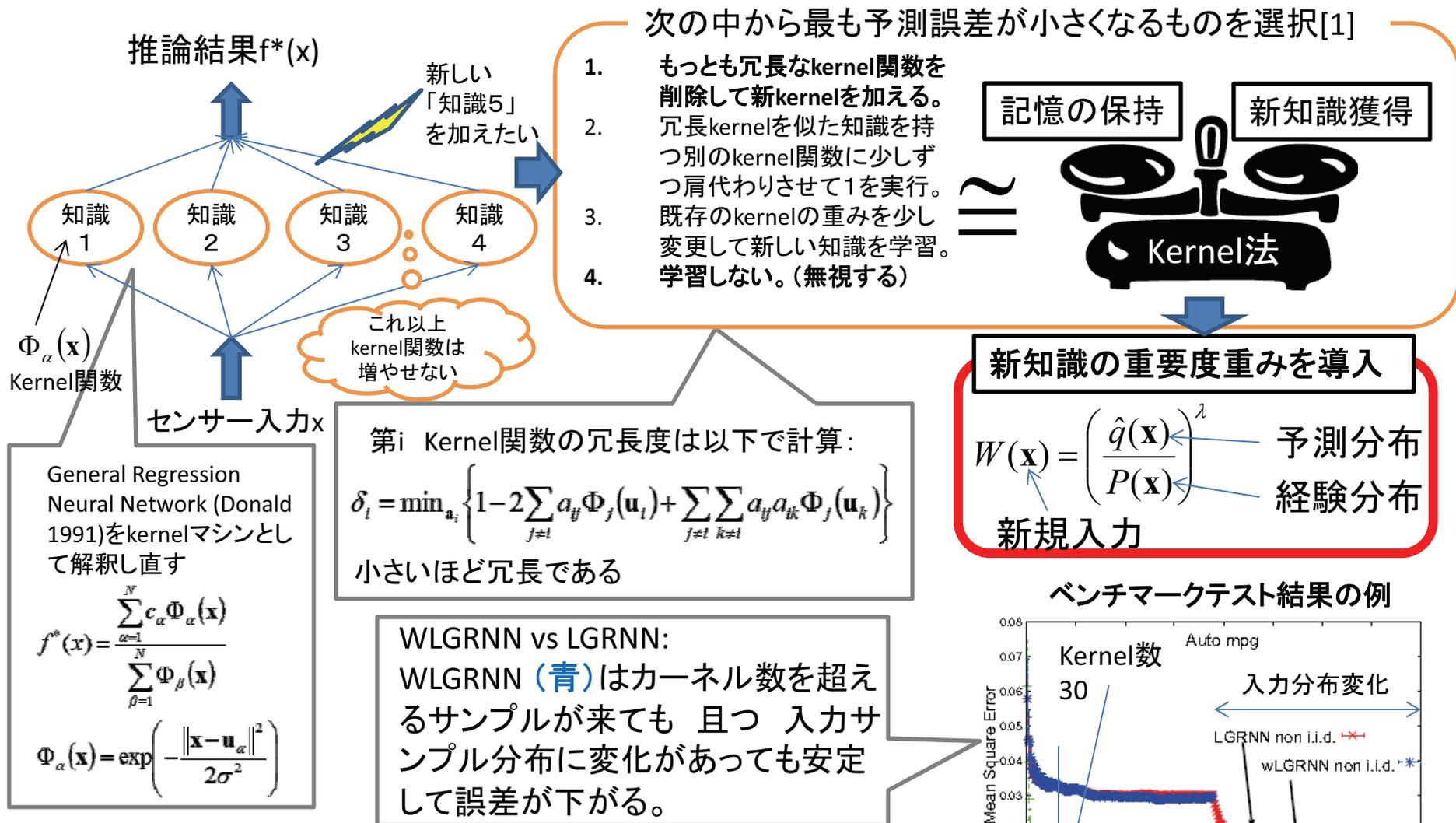


学習機能付き高速
MPPTコンバータ

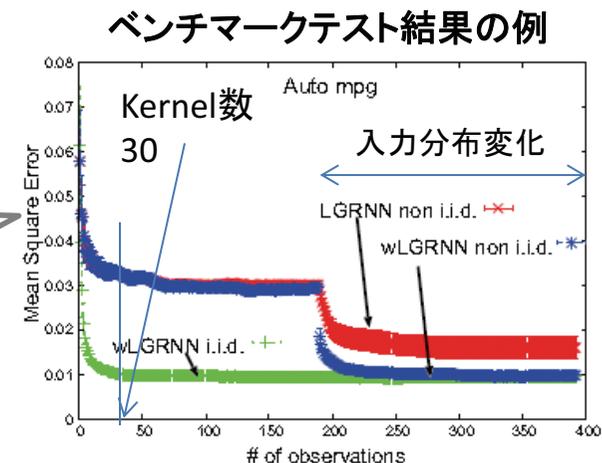
要件:

- 厳しいリソース制限があっても良い。
- 少ない計算量で実用的速度で計算できる。
- 非i.i.d.サンプル(非独立、経年変化するデータ)でも安定した学習結果が得られる。
- **重要サンプルは漏らさず学習し、不要サンプルは忘却する。**

Importance Weighted Limited General Regression Neural Network (WLGRNN)



[1] Koichiro Yamauchi . "Pruning with Replacement and Automatic Distance Metric Detection in Limited General Regression Neural Networks", IJCNN, pp. 899--906, July, (2011)



D-142 関数データに対する回帰分析における 収束レート

綾野孝則 (大阪大学)
t-ayano@cr.math.sci.osaka-u.ac.jp

X : 説明変数、 Y : 目的変数

従来 \dots $X \in \mathbb{R}^d$ 、 $Y \in \mathbb{R}$

今回 \dots $X \in$ 関数空間、 $Y \in \mathbb{R}$ (関数データ解析)

\mathcal{D}^m : m 回微分可能な関数全体

$W^{s,p}$: Sobolev 空間、 B_{pq}^s : Besov 空間

\mathcal{H}_K : 再生核ヒルベルト空間

- Biau et al. 2010

X が上記の関数空間に含まれるときの k -NN 推定量における収束レートの上界を導出。 $O\left((\log n)^{-2m/d}\right)$

- 本研究

- kernel 推定量でも同様のレートが達成出来る。

- 回帰関数の滑らかさ γ を導入。 $O\left((\log n)^{-2m\gamma/d}\right)$

本研究の目標 …… 上記のレートが minimax レートを達成していることを示すこと。

参考文献

G.Biau, F.Cérou, A.Guyader (2010), Rates of Convergence of the Functional k -Nearest Neighbor Estimate, IEEE Trans.Inform.Theory 56,No.4,2034-2040.

D-143: 1クラスラベルに注目したt統計量とそのLassoへの応用

小森 理, 江口 真透
統計数理研究所

判別関数 $F(\mathbf{x}) = \beta^\top \mathbf{x}$ と生成関数 $U: \mathbf{R} \rightarrow \mathbf{R}$ を用いて, 観測データ $\{(x_i, y_i) : i = 1, \dots, n\}$ に対し, 次のような統計量を考える.

t_U 統計量

$$\bar{t}_U(\beta) = \frac{1}{n_1} \sum_{y_i=1} U\left(\frac{\beta^\top (x_i - \bar{x}_0)}{\sqrt{\beta^\top S_0 \beta}}\right),$$

- $n_y = \#\{i | y_i = y\}$, $\bar{x}_y = \frac{1}{n_y} \sum_{y_i=y} x_i$
- $S_0 = \frac{1}{n_0} \sum_{y_i=0} (x_i - \bar{x}_0)(x_i - \bar{x}_0)^\top$
- $\bar{t}_U(a\beta) = \bar{t}_U(\beta)$ ($\forall a > 0$)
- $U(w) = w$ のとき, t_U 統計量は通常の t 統計量に帰着.

ベイズルール

$F^*(x)$ と $P(Y = 1|X = x)$ が一対一の関係を持つための必要十分条件は

結果1

$$U'(w) = c, \text{ or } U'(w) - wU''(w) > 0, \forall w \in \mathbf{R}.$$

である。ただし, c は任意の定数。

- $U(w) = w$: 線形関数
- $U(w) = \Phi(w)$: 正規分布の標準分布関数

漸近分散

結果2

$$\frac{\{\beta^{*\top}(\mu_1 - \mu_0)\}^2}{\beta^{*\top}\Sigma\beta} \Sigma^* \mathbb{V}_{\text{asy}}(\widehat{\beta}) \Sigma^*$$
$$= \frac{\mathbb{E}\{U'(W)^2|Y=1\} + \sqrt{\frac{\pi_1}{\pi_0}} \left[\mathbb{E}\{U'(W)W|Y=1\} \right]^2 + \sqrt{\frac{\pi_1}{\pi_0}} \left[\mathbb{E}\{U'(W)|Y=1\} \right]^2}{\left[\mathbb{E}\{U'(W)|Y=1\} \right]^2} \Sigma^*$$

よって漸近分散を最小にする $U_0(w)$ は

$$U_0(w) = -\frac{1}{2} \frac{\sqrt{\frac{\pi_1}{\pi_0}} \mathbb{E}(W|Y=1) \mathbb{V}(W|Y=1)}{1 + \sqrt{\frac{\pi_1}{\pi_0}} \mathbb{V}(W|Y=1)} \left[w - \frac{1 + \sqrt{\frac{\pi_1}{\pi_0}} \mathbb{E}(W^2|Y=1) \mathbb{V}(W|Y=1)}{\sqrt{\frac{\pi_1}{\pi_0}} \mathbb{E}(W|Y=1) \mathbb{V}(W|Y=1)} \right]^2.$$

動的サブセットクラスタリング

石黒 勝彦、上田 修功、澤田 宏 (NTT)

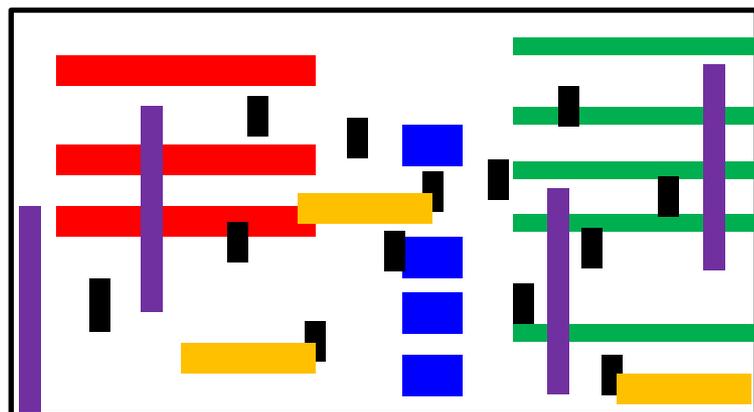
目的: 時系列データから、**時間特異的に重要な部分だけ**をクラスタリングしたい

例えば:

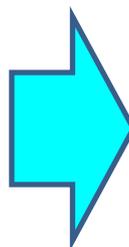
ある期間で特徴的な購買パターン(ブーム)を抽出

大量のセンサーデータから、役立つ部分だけをまとめて発見

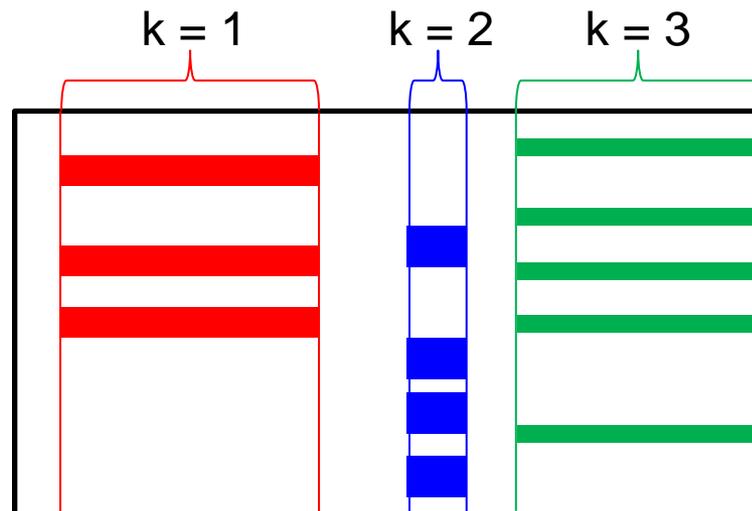
オブジェクト(特徴)



時刻 t

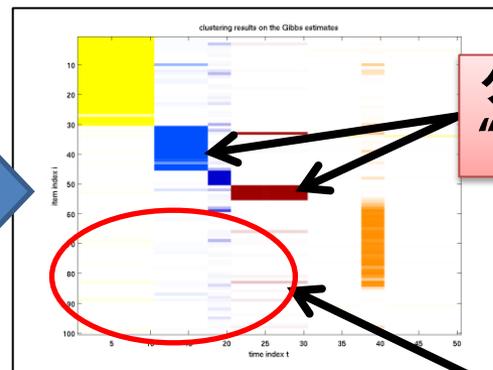
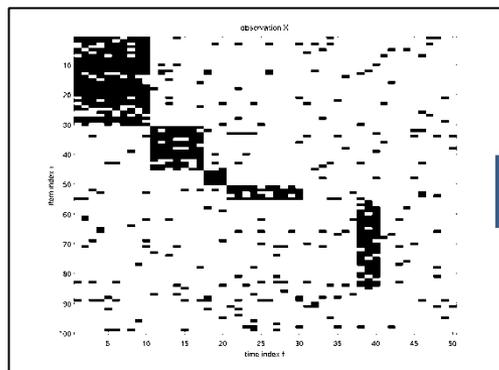


オブジェクト(特徴)

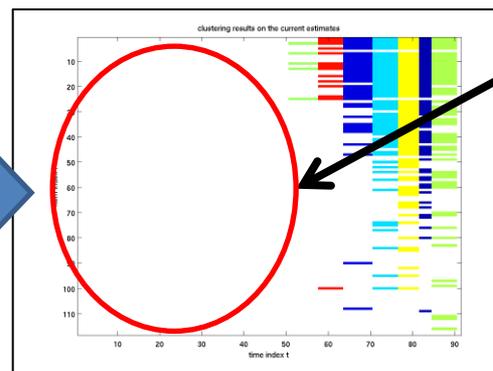
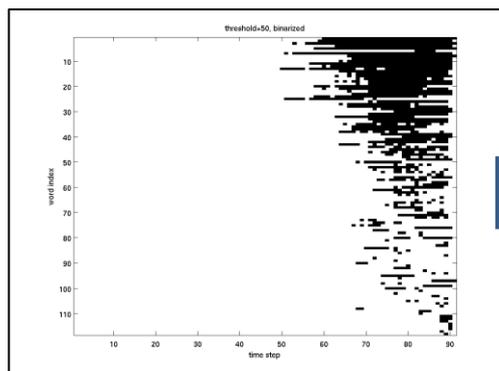


時刻 t

方法:時系列データをHMMでモデル化する際に、
 クラスタ=stateの区別に役立つ特徴の選択も行う



クラスタごとに異なる
 “重要”特徴を抽出



役に立たない
 特徴は無視して
 クラスタリング



クラスタごとの“重要”特徴を自動的に抽出
 クラスタ数もノンパラメトリックベイズで自動的に推定

D-145 ロバスト最適化による判別モデル

武田朗子⁽¹⁾ 参木裕之⁽¹⁾ 金森敬文⁽²⁾

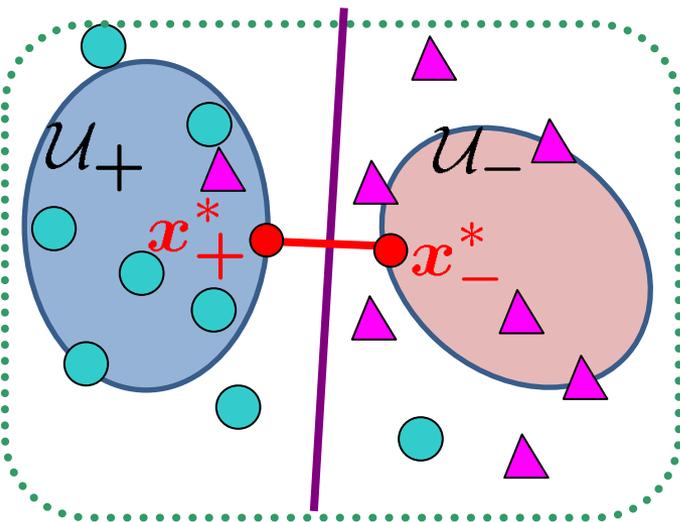
⁽¹⁾慶應義塾大学

⁽²⁾名古屋大学

二値判別問題の観測データには、観測誤差等による不確実性がある

→ ロバスト最適化法による、データの不確実性に頑健な判別モデル

$$\max_{\|\mathbf{w}\|^2=1} \min_{\mathbf{x}_+ \in \mathcal{U}_+, \mathbf{x}_- \in \mathcal{U}_-} (\mathbf{x}_+ - \mathbf{x}_-)^T \mathbf{w}$$



- ✓ 各クラスの観測点の平均 (\mathbf{x}_+ , \mathbf{x}_-) を不確実な入力
- ✓ \mathbf{x}_+ , \mathbf{x}_- の動きうる範囲 (\mathcal{U}_+ , \mathcal{U}_-) を設定
- ✓ 最悪な \mathbf{x}_+^* , \mathbf{x}_-^* を想定して最適化

Robust Classification Model (RCM)

$$\max_{\|\mathbf{w}\|^2=1} \min_{\mathbf{x}_+ \in \mathcal{U}_+, \mathbf{x}_- \in \mathcal{U}_-} (\mathbf{x}_+ - \mathbf{x}_-)^{\top} \mathbf{w}$$

✓ $\mathcal{U}_+, \mathcal{U}_-$ の与え方により、SVMや minimax probability machine, フィッシャー判別モデルと等価

✓ RCMは確率分布の不確実性を考慮した期待損失最小化:

$$\min_{\mathbf{w}: \|\mathbf{w}\|^2=1, b} \max_{\substack{p(\mathbf{x}|+1) \in \mathcal{P}_+ \\ p(\mathbf{x}|-1) \in \mathcal{P}_-}} \mathbb{E}[\ell(y(\mathbf{x}^{\top} \mathbf{w} + b))]$$

をよく近似

✓ $\mathcal{U}_+, \mathcal{U}_-$ に交わりがないと凸最適化問題に変形可能、交わりがあると本質的に非凸最適化問題

→ 局所最適解法の提案と、よい数値実験結果の報告

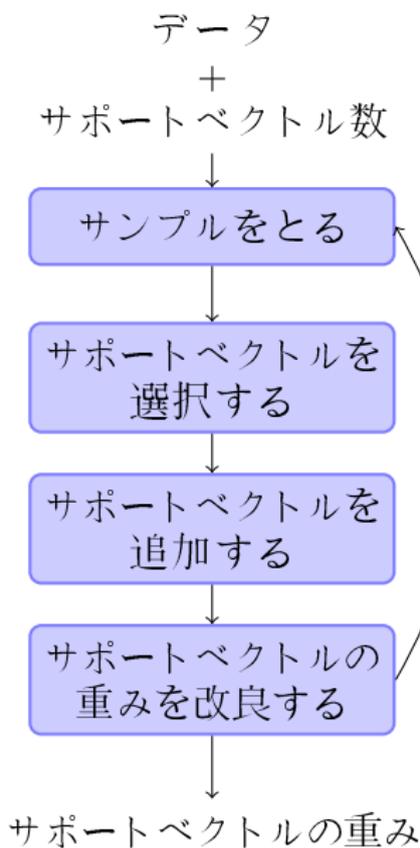
D-146 頑健なスパースカーネル 分類器の学習

Mathieu Blondel

神戸大学大学院システム情報学研究科

2011-11-10

キーワード：カーネル，SVM，パーセプトロン



- スパースカーネル分類器を学習するための最適化アルゴリズムを提案
- 主問題によるアルゴリズム
- ラベルノイズに対して頑健である新たなサポートベクトル選択方法を利用
- 任意の損失関数ができる
- 訓練データを効率的に活用する