

T-17 Regularized Policy Gradients: Direct variance reduction in policy gradient estimation

Tingting Zhao¹, **Gang Niu**², Ning Xie³, Jucheng Yang¹, Masashi Sugiyama²



TUST



UTokyo



TJU

- Reinforcement Learning :: Policy Search :: Policy Gradients
- Further stabilize the state of the art by *directly regularizing* the variance of the gradient estimates

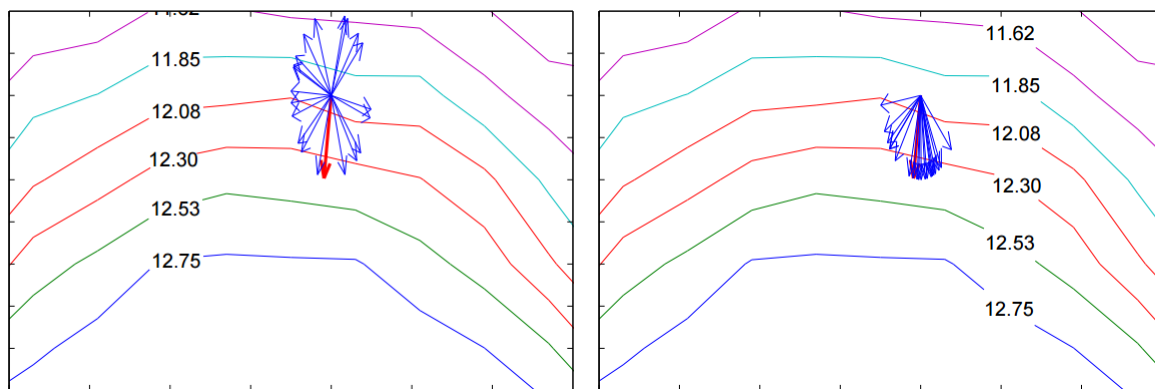
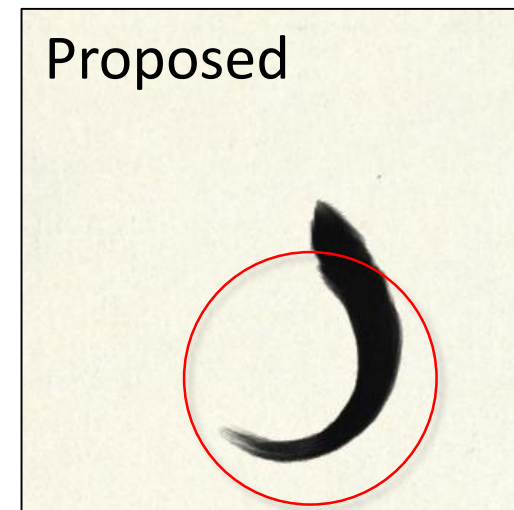
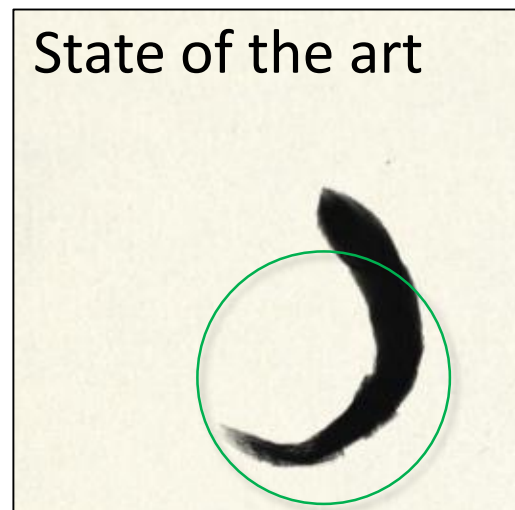


Illustration of large Var vs. small Var
(Gradient of objective; Estimate of gradient)

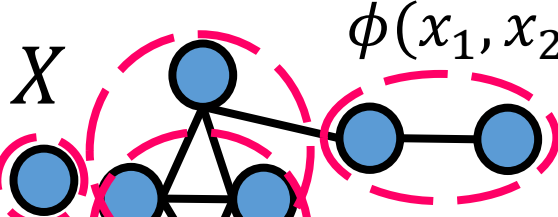


Applied to stroke-based rendering

T-22 Learning Structure of Partial Markov Random Field via Partitioned Ratio

Song Liu (ISM), Taiji Suzuki (Tokyo Tech), Masashi Sugiyama (UTokyo), Kenji Fukumizu (ISM)

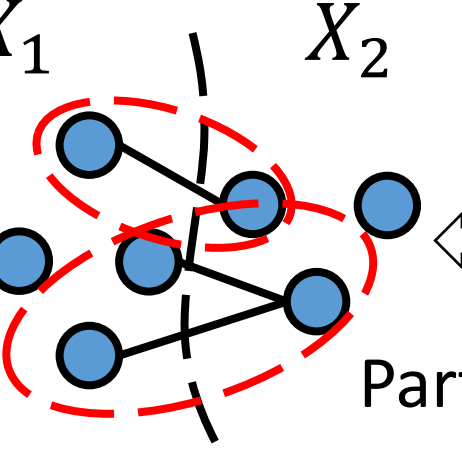
Target: learning **sparse** and **partial** connectivity of an MRF.



X $\phi(x_1, x_2)$ $\iff p(\mathbf{x}) \propto \prod_{c \in C^*} \phi(\mathbf{x}_c)$

Hammersley Clifford Theorem

Joint distribution factorizes over its cliques.



X_1 X_2 $\iff \frac{p(x_1, x_2)}{p(x_1)p(x_2)} \propto \prod_{b \in B^*} \phi(\mathbf{x}_b)$

Partitioned Ratio

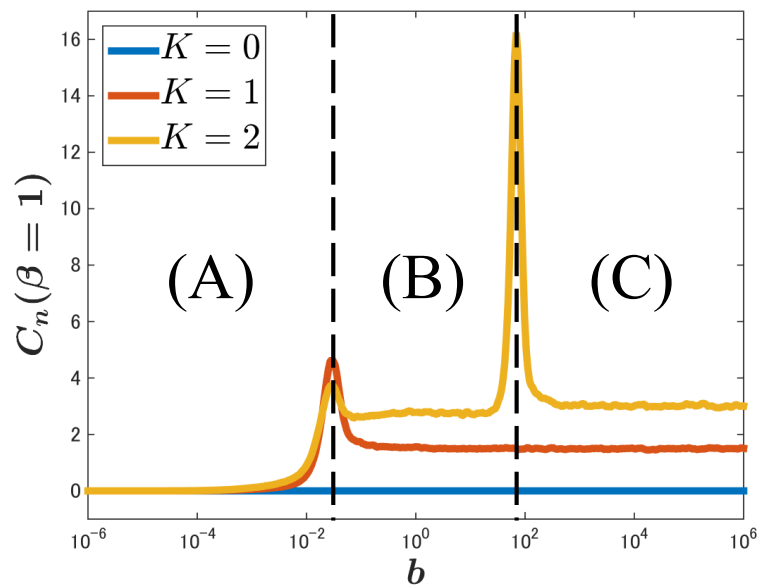
Partitioned Ratio factorizes over its **bridges**.

Contribution: A support consistent method is proposed and analyzed.

T-24 比熱から見るベイズ推定における必要最小データ量

○徳田悟, 永田賢二, 岡田真人 (東大)

学生優秀プレゼンテーション賞対象



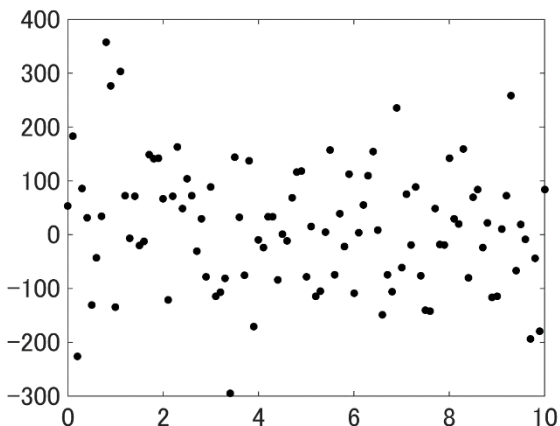
目的

統計的推測の状態変化を説明する理論を構築し、「必要」なサンプル数やノイズレベルを見積もりたい。

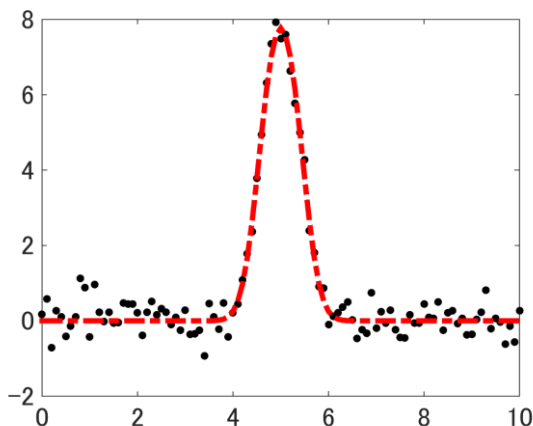
方法

比熱という新たな統計量を導入。
ベイズ統計の漸近理論を有限拡張。

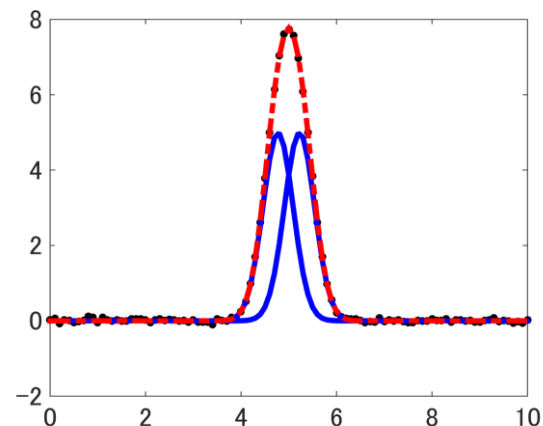
(A) 高温相



(B) 中温相

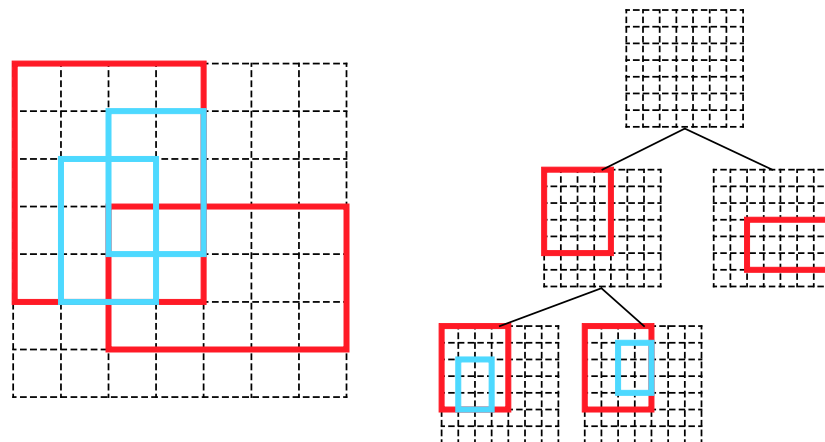


(C) 低温相



概要

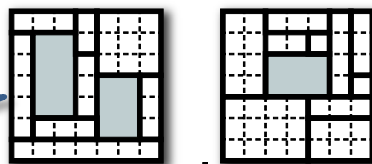
無限サイズの行列に
 対する任意のR木(領域の
 包含を親子関係とする木構造)
 を表す確率過程の構成



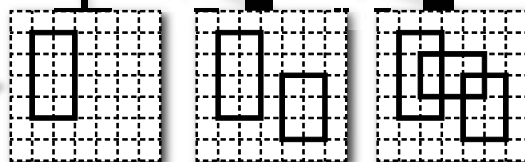
確率過程構成

二重マルコフ過程による無限再帰

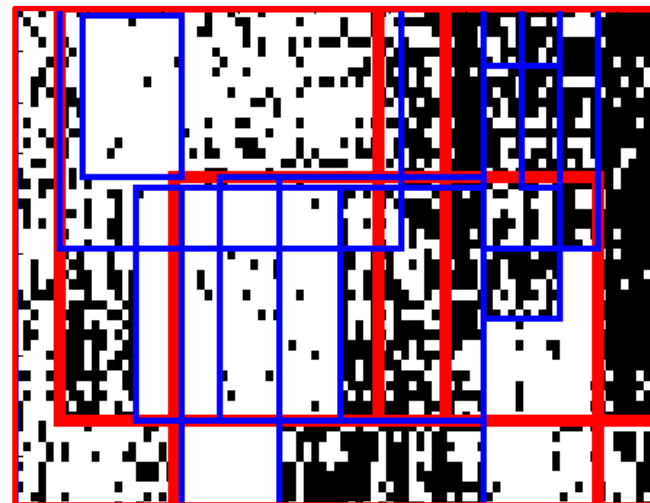
領域候補の生成



領域の抽出



応用: 関係データ解析



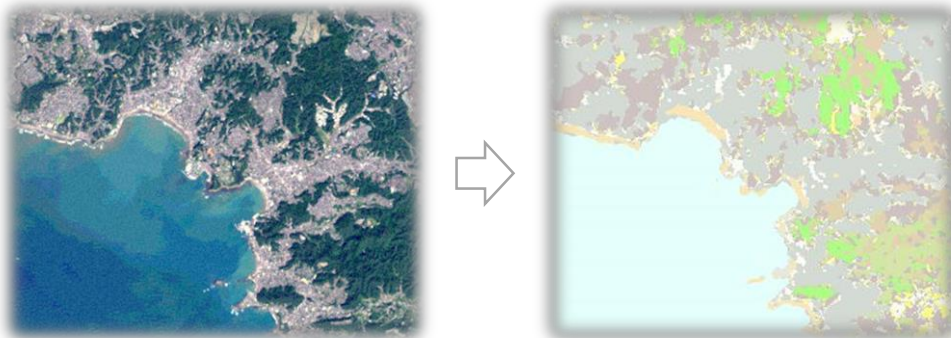
少数の教師データと多クラス問題のための再帰的アンサンブル土地被覆分類

大屋優・金盛克俊・大和田勇人（理科大）

目的 放置林のように教師データの不足する場所に対しても詳細な土地利用マップを作成する

方法 異なる時刻データからのアンサンブルと教師データからの周辺への影響度を考慮して半教師データを再帰的に求める

結果 わずかな教師データと特徴表現の乏しいデータ（マルチスペクトルデータ）からでも詳細（多クラス）な土地の分類が行えた



T-29 Weighted Semi-Binary NMFに基づく機器別消費電力推定

松本眞子・藤本悠・林泰弘(早大)

※学生優秀プレゼンテーション賞対象

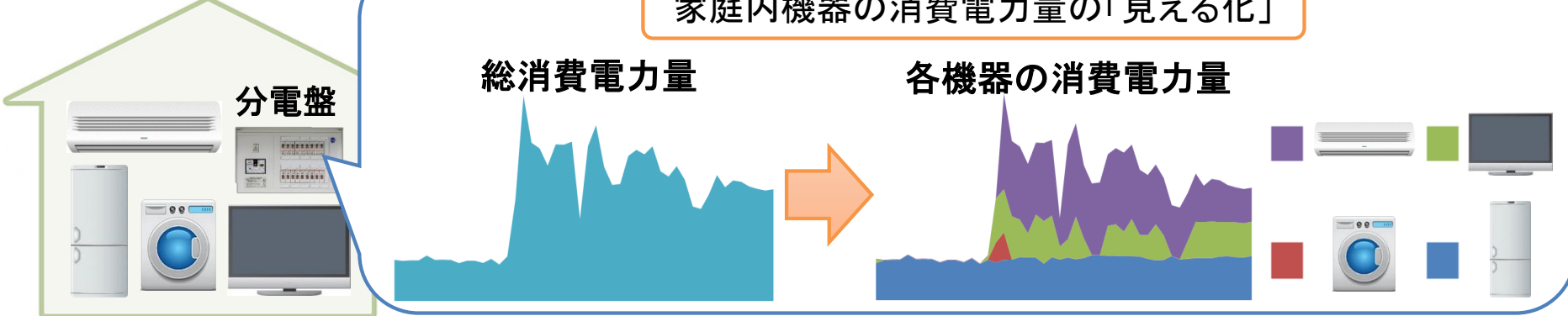
◆背景と目的

家庭内機器の消費電力量の「見える化」

総消費電力量

各機器の消費電力量

分電盤



◆機器別消費電力推定の枠組み (Semi-Binary NMF)

総消費電力量
 Y

\cong

機器使用パターン(基底)
 X

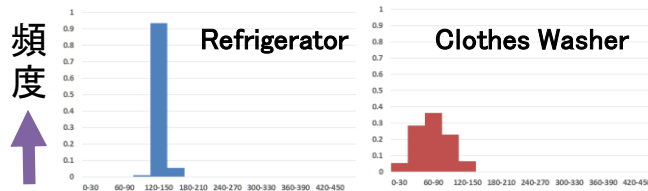
\times

基底の稼働状況 (ON or OFF)
 A

◆提案手法

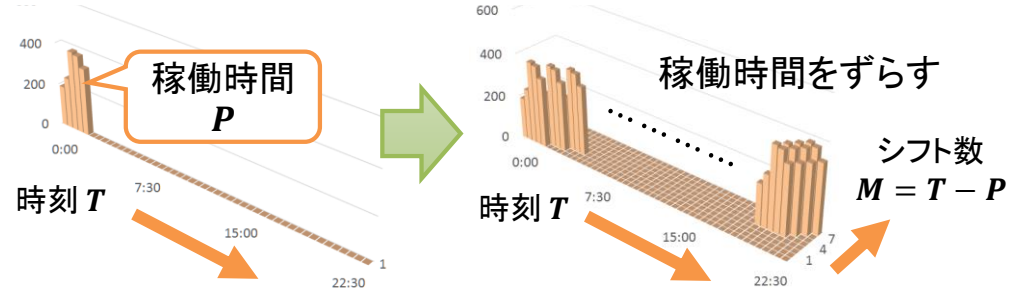
- 機器の消費電力量にはばらつきがある
- 稼働時間が短く、稼働時間が日毎に異なる機器が存在

機器稼働時の消費電力量の分布

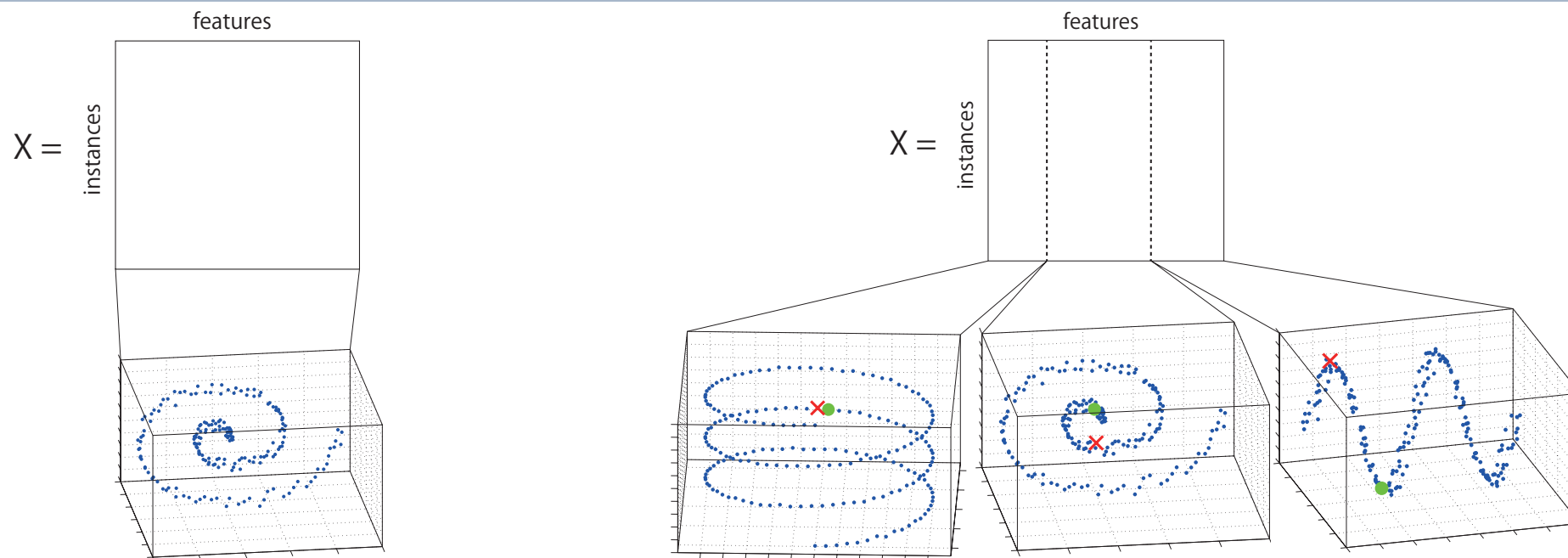


消費電力量 [Wh/30min]

機器毎の消費電力量のばらつき
を考慮した学習



稼働時間を短くし、時間帯シフト
をうまく表現する基底を学習



標準的な多様体学習：特徴空間全体が多様体構造を共有

本研究での問題設定：複数の特徴グループが各々異なる多様体に従う（異なる近傍関係を持つ）

全ての特徴量が同じ多様体に従うという多様体学習の仮定は非常に強く，複数の多様体によって構成されるようなより一般的な特徴空間を扱うことはできない

本研究では局所線形モデルと L1 グラフを利用して，特徴の**グループとグループ毎の近傍関係を同時に推定**する

- 高次元データの潜在的な構造を複数の多様体で記述し，解析できる
- 非線形の特徴クラスタリングとしての解釈が可能

T-31 能力テスト得点の非負行列分解

兼村厚範^{*}、大成弘子[†]、鹿内学[†]、橋本将崇[†]、赤穂昭太郎^{*}

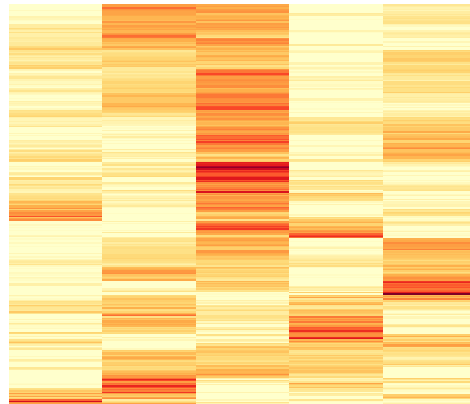
^{*} 産総研 [†] リクルートキャリア

項目応答パターン (300 x 58)

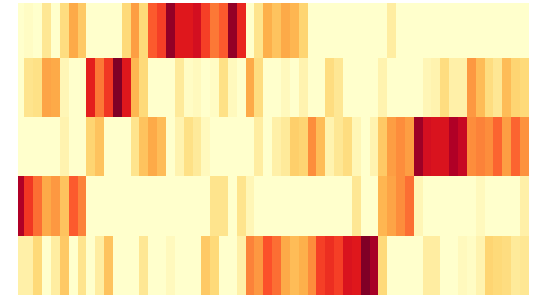
0	1	...
1	0	...
1	1	...
...

=

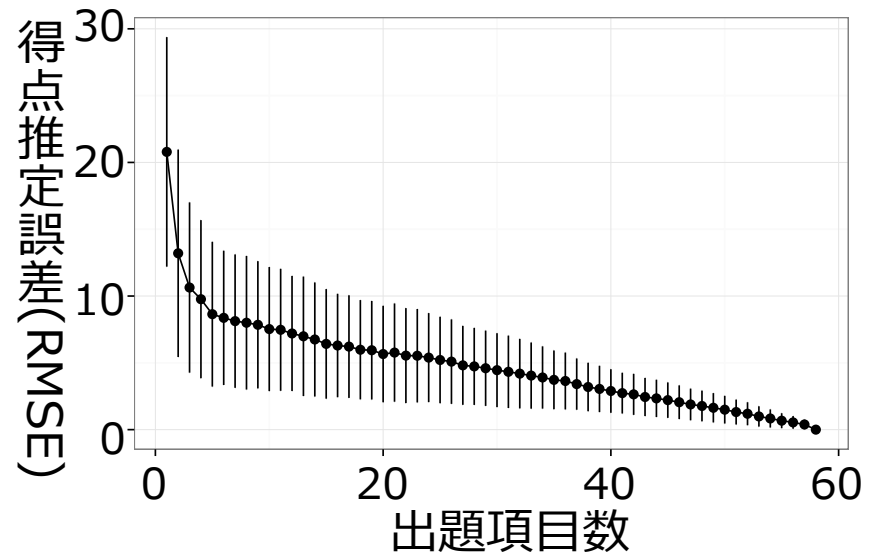
係数 (300 x 5)



スキルセット (5 x 58)

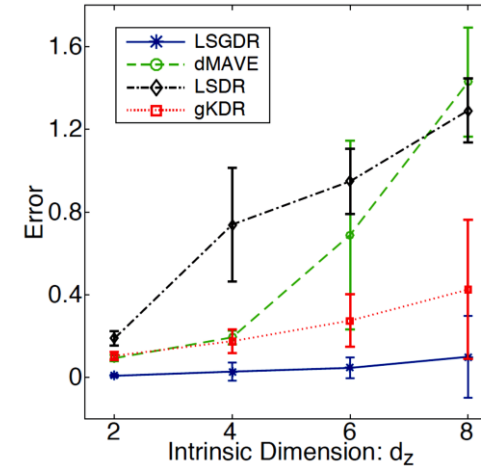
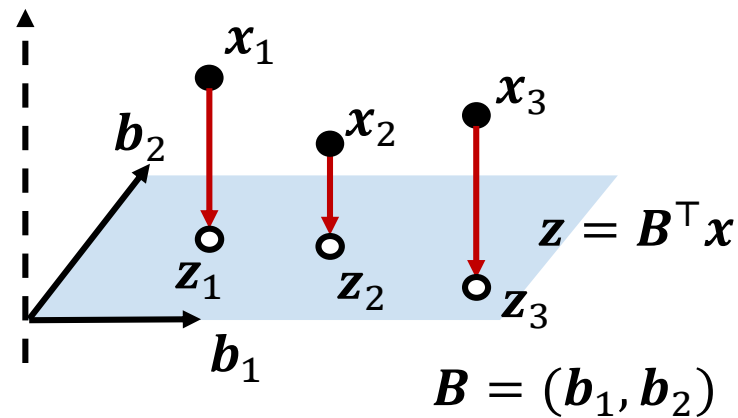


- 項目回答型の能力テストやその設計に要する問題数の削減
 - テスト時間短縮
 - 設計者の負担低下
- ⇒ 非負行列分解による欠測推定



T-33 Sufficient Dimension Reduction via Direct Estimation of the Gradients of Logarithmic Conditional Densities

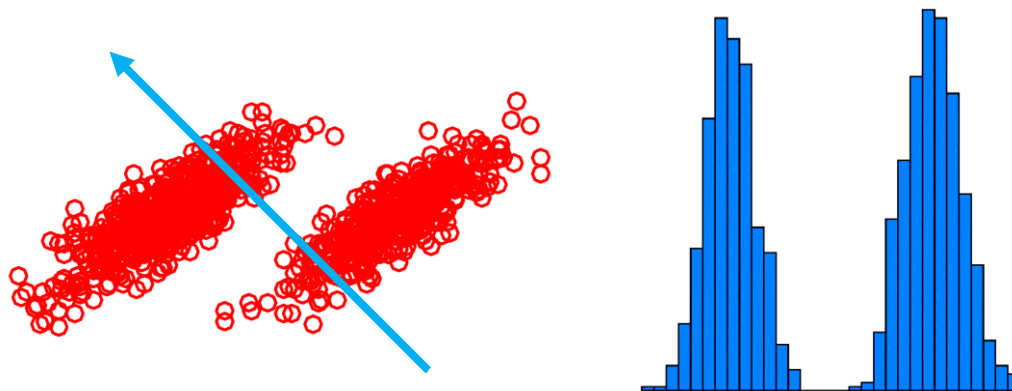
Hiroaki Sasaki (UTokyo), Voot Tangkaratt (UTokyo) & Masashi Sugiyama (UTokyo)



- **Problem:** Supervised linear dimension reduction.
 - Obtain the transformation matrix B such that $p(y|x) = p(y|z = B^T x)$.
- **Approach:** Apply the eigenvalue decomposition to $E\{\nabla_x \log p(y|x) \nabla_x \log p(y|x)^T\}$.
 - B is estimated in a closed form.
- **Our Contribution:** A novel direct estimator for $\nabla_x \log p(y|x)$.
- **Our Method:** *Least-Squares Gradients for Dimension Reduction (LSGDR)*.
 - An accurate and computationally efficient method.

Non-Gaussian Component Analysis with Log-Density Gradient Estimation

Hiroaki Sasaki (UTokyo), Gang Niu (UTokyo) & Masashi Sugiyama (UTokyo)

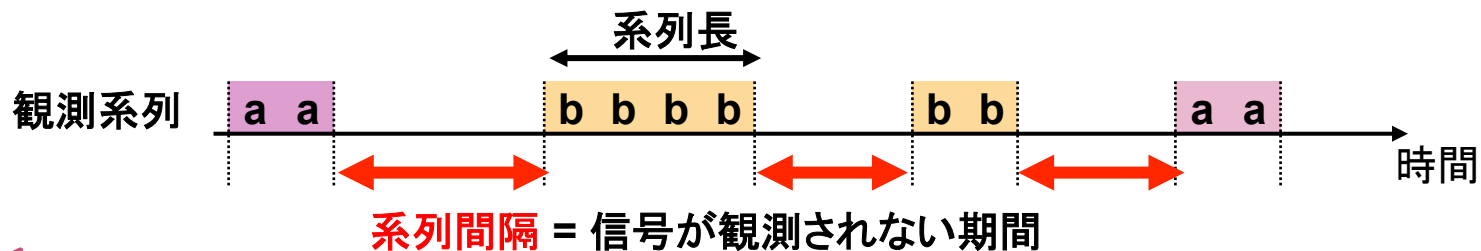


- **Non-Gaussian Component Analysis (NGCA)** [Blanchard et al., 2006]: An *unsupervised* framework to find a linear non-Gaussian subspace.
- **Approach:** Eigenvalue decomposition to $E\{(\nabla_x \log p(\mathbf{x}) + \mathbf{x})(\nabla_x \log p(\mathbf{x}) + \mathbf{x})^\top\}$.
 - The non-Gaussian subspace is identified **in a closed form**.
- **Ingredient:** Direct estimator for $\nabla_x \log p(\mathbf{x})$ *without density estimation*.
 - **Least-squares log-density gradients** [Cox, 1985; Sasaki et al., 2014].
- **Results:**
 - Convergence rate analysis to the true subspace.
 - Experimental demonstration of the superior performance.

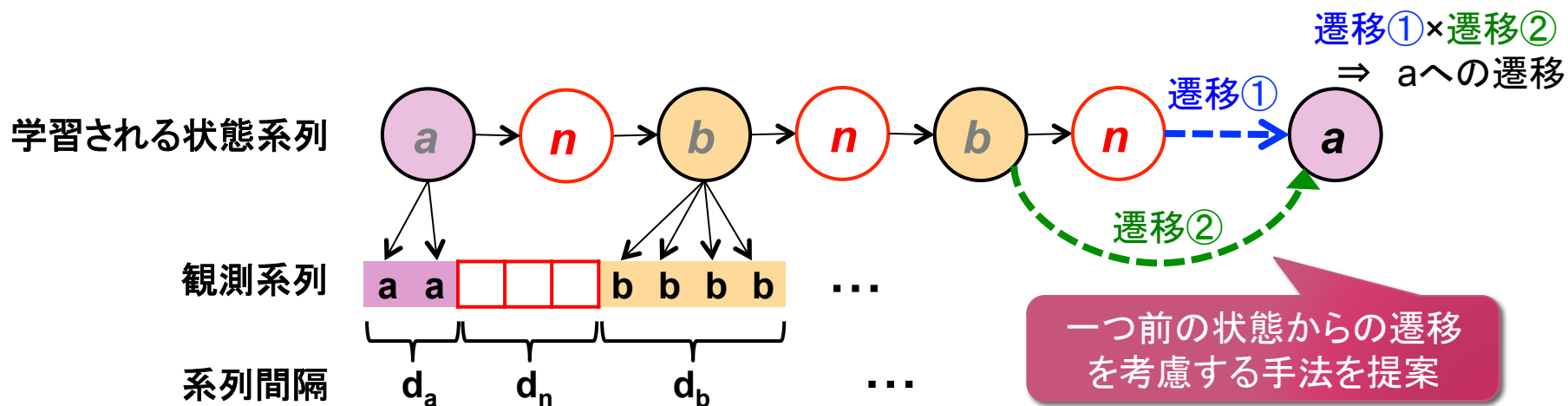
【T-36】 ○成松宏美, 笠井裕之(電通大)

系列間隔状態ノードを有する隠れセミマルコフモデルの評価

系列長を表現可能なHSMMを用いて系列間隔をモデル化する



課題

HSMMの1状態として, “系列間隔状態 n ” を表現すると...本来の状態系列 $\text{a} \rightarrow \text{b} \rightarrow \text{b} \rightarrow \text{a}$ の表現性能が劣化

T-37: Performance degradation of AMP for Ising perceptron when the system size is small 栗谷 亜理世 (京都大学システム科学専攻)

近似メッセージ伝搬方 (AMP) は問題サイズが大きい場合には漸近的に最適解を達成する

一方で、問題サイズが比較的小さい場合、性能は低下するがその具体的理由は分からない

本研究では AMP を、その土台となるアルゴリズム (BP, ガウス近似のみの AMP) と比較する事で性能低下のメカニズム、近似による影響を探る

学生優秀プレゼンテーション賞対象

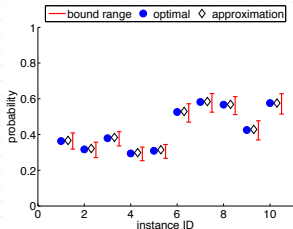
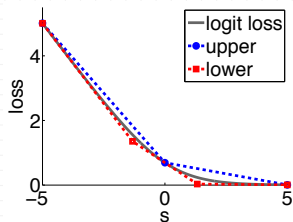
準同型暗号を用いた

経験損失最小化のための秘密精度保証

高田 敏行 †, 花田 博幸 †, 佐久間 淳 ††, 竹内 一郎 †

† 名古屋工業大学, †† 筑波大学

- ▶ プライバシ保護経験損失最小化
 - ▶ 非線形関数に近似関数を用いる必要あり
 - ▶ 得られる解と最適解に誤差が生じる
- ▶ 提案法
 - ▶ プライバシ保護を達成しながら最適解との差を保証
 - ▶ 区分線形関数を用いて元の損失関数を上下からバウンド



T-39 分布の局所平滑化による正則化の提案

宮戸岳 前田新一 小山雅典 中江健 石井信(京大)

- 回帰,分類タスクについて

我々がよく扱う自然画像や時系列データには
「入力空間上の滑らかさ」の性質が存在



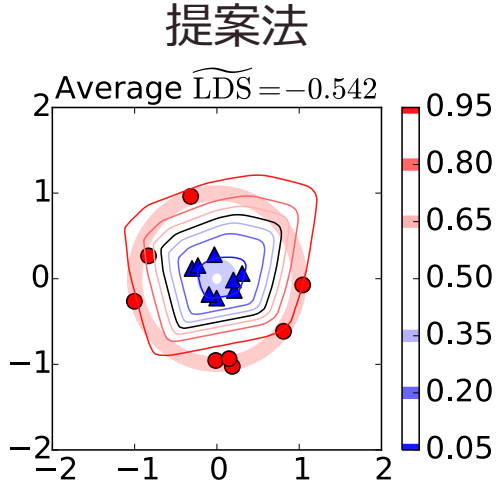
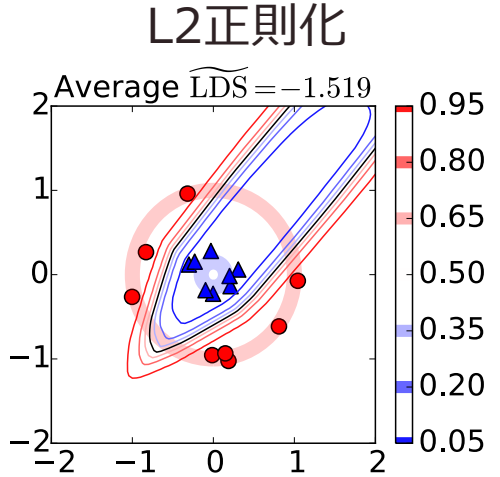
モデル $p(y|x, \theta)$ に関して、入力 x に対して
出力 y の分布が滑らかに変化するべき

- これを実現する正則化関数として、モデル分布の入力データ点周りの局所的な滑らかさ (Local Distributional Smoothness, LDS)を導入:

$$LDS(x, \theta) \equiv -KL[p(y|x, \theta) || p(y|x + r_{v\text{-adv}}, \theta)]$$

$r_{v\text{-adv}}$: ある恣意的な擾乱

- LDSによる正則化は、ベンチマークデータに対して最新的手法とほぼ同等の性能を達成!



T-40:

ベイジアンマスキングを用いたスパース特徴選択

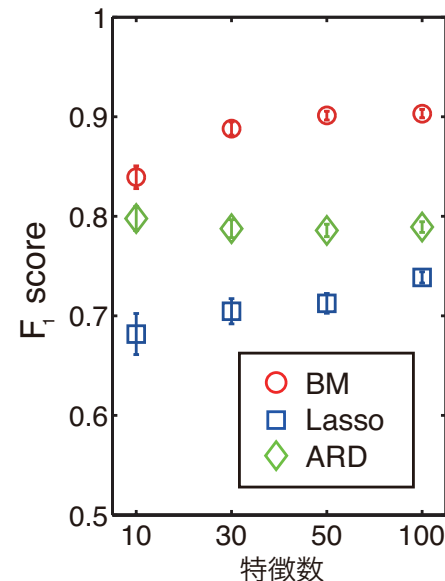
近藤 洋平 (京大), 林 浩平 (NII), 前田 新一 (京大)

目的: 線形回帰 $y = X\beta + \varepsilon$ におけるスパース特徴選択

提案: 入力をマスクする変数 Z を $y = (Z \circ X)\beta + \varepsilon$ と導入し
情報量基準FICに基づくベイジアンモデル選択を行う

検証: 数値実験にてLassoとARDを
大きく超える F_1 スコアを達成

$$F_1 = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$



ガウシアンプロセスカーネル法による 非線形テンソル学習およびマルチタスク学習への応用

金川平志郎¹・鈴木大慈^{1,2}

¹東京工業大学 ²さきがけ, 科学技術振興機構

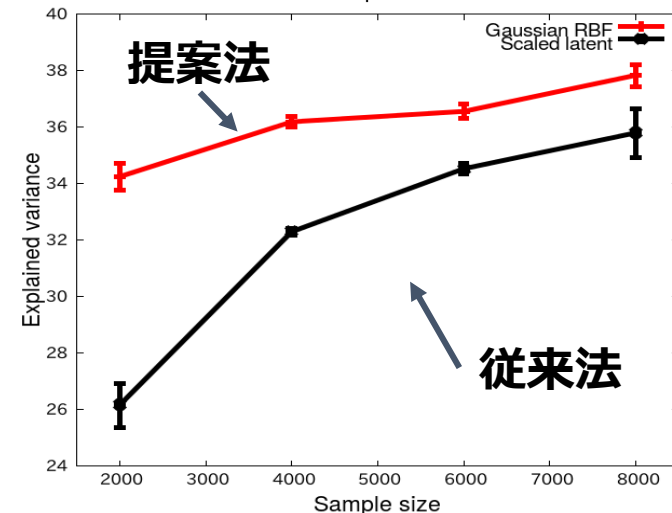
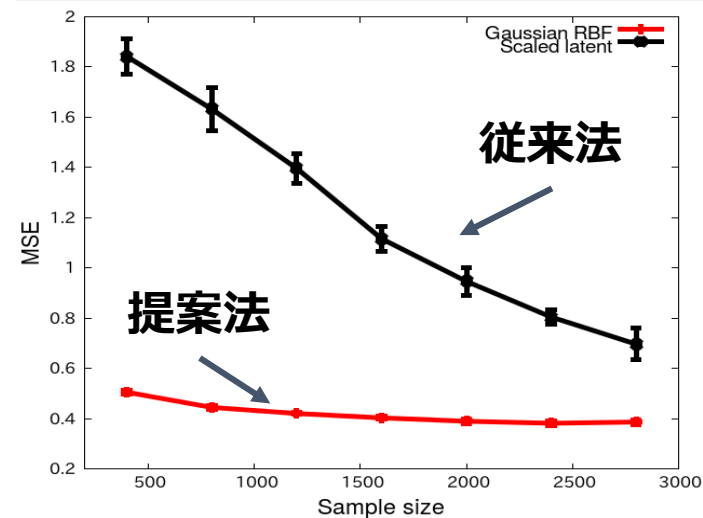
学生優秀プレゼンテーション賞対象

- We propose a **nonparametric Bayesian tensor learning method**:

$$y = \sum_{r=1}^{d^*} \prod_{k=1}^K f_r^{(k)}(x^{(k)}) + \epsilon$$

Gaussian process prior is employed for $f_r^{(k)}$

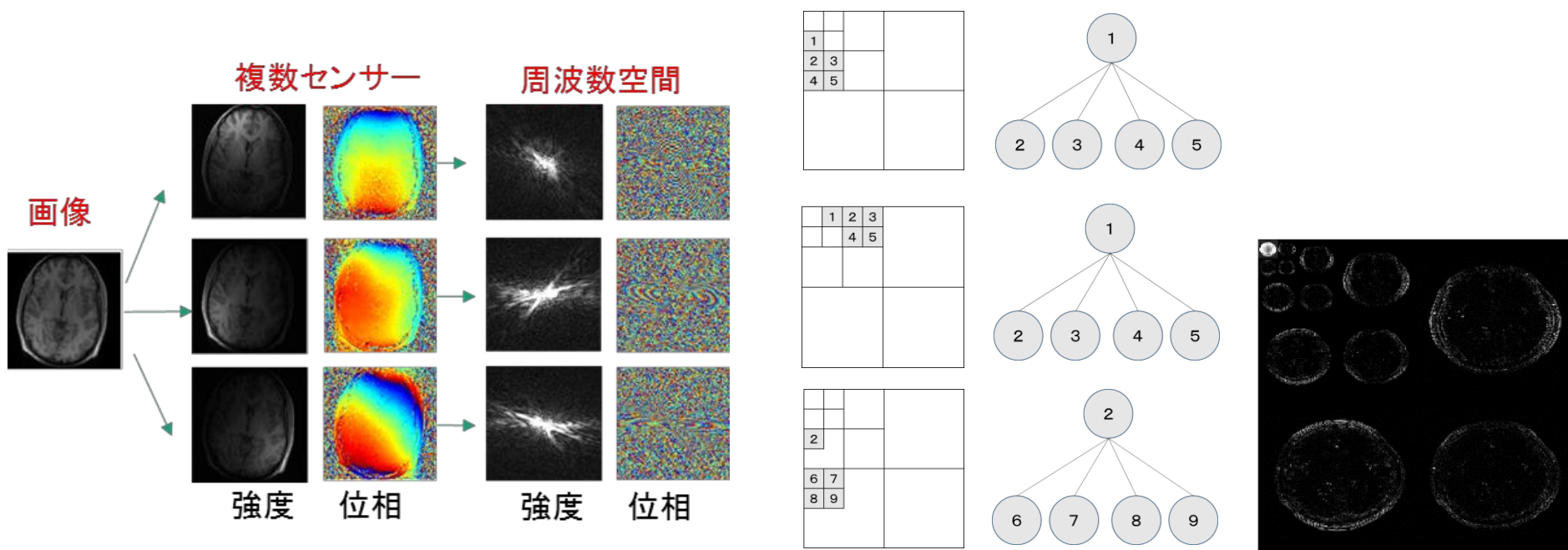
- Fast convergence rate is achieved **without assuming any strong convexity on the design**
- Application to multitask learning
 - Proposed method outperforms existing linear methods in experiments on real world data sets



SENSE法とwavelet木構造の疎性を用いた複数コイルMR画像再構成法

笠原勇布・井上真郷(早大)・富樫かおり(京大)

- 複数コイルMR画像再構成に構造を用いた正則化
- wavelet木構造の疎性を利用



- 単純なwavelet変換の疎性よりも性能の改善が見られた

T-43 松田源立(東京大学)

Adaptive Objective Function of ICA by Gaussian Approximation in Second-Order Polynomial Feature Space

- ◆ 概要：独立成分分析(ICA)の適応的な目的関数の構築
- ◆ ICA：元信号の非ガウス性を利用して、観測信号のみから元信号を推定する手法
- ◆ 従来のICA：元信号の確率分布を決め打ちまたは切換え型で付与
- ◆ 提案するICA：元信号分布の非ガウス性の程度を適応的かつ滑らかに推定

順序回帰問題に対するL1正則化の導入

T-44

長島主尚, 井上真郷 (早大)

順序データ

- 服のサイズ
- 学校の成績
- レビューの評価値

+

L1正則化

$$\lambda \sum_{m=1}^M |w_m|$$

FISTA

- スパースな解
- 既存手法と同等の汎化性能
- 大規模データで高速な可能性

統一的2値判別モデルに対する効率的な加速近接勾配法

伊藤直紀(東大) ・ 武田朗子 (東大) ・ TOH Kim-Chuan (NUS)

- **従来:** 2値判別モデルの実用的な解法 … 各モデルの構造に**特化**

2値判別モデル	定式化	解法
サポートベクターマシン	ヒンジロス最小化	SMO, 座標降下法 etc.
ミニマックス確率マシン	誤判別確率最小化	内点法
フィッシャーの判別分析	分散比最大化	固有値分解

- 実はノルム最小化問題として統一的に定式化できる (先行研究)

$$\text{統一的2値判別モデル} \quad \min \{ \|Aq - b\|_2 \mid q \in Q \}$$

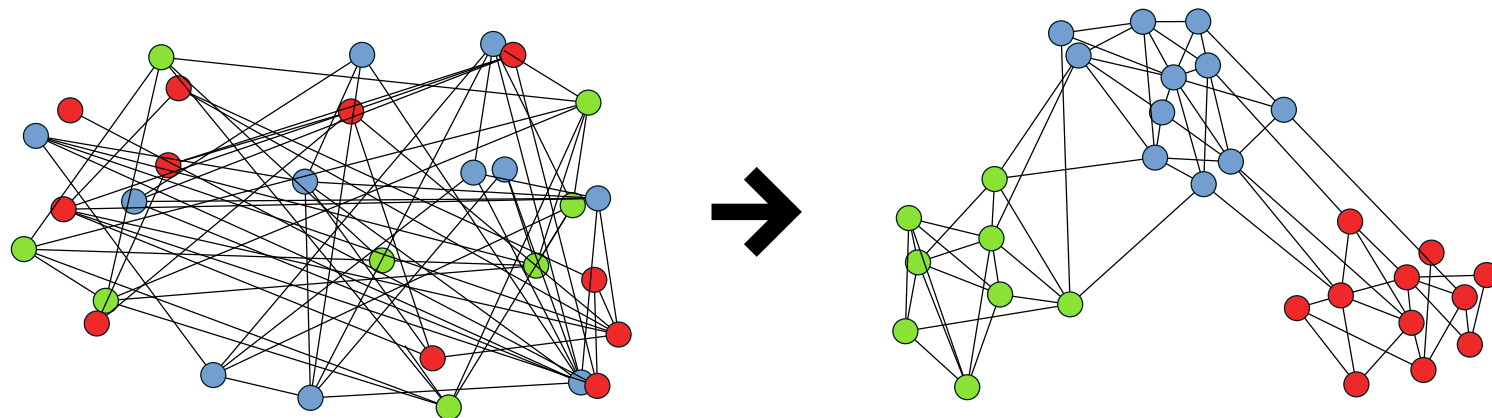
- **本研究:** 共通の問題構造を利用した**汎用的解法**を提案

- 射影アルゴリズムを設計
- 加速近接勾配法 + 高速化の工夫
- 例えばSMO, 座標降下法より高速

モデル選択が柔軟に
できるようになると
期待される

D-02: 確率ブロックモデルの漸近近似にもとづく トラクタブルなベイズ推論およびモデル選択

林浩平(NII), 小西卓也(NII), 川本達郎(東工大)



- 確率ブロックモデル(SBM)の近似推論
 - 周辺化を組み込んだBP則
 - クラスタ数を自動決定するアルゴリズム
 - ノンパラベイズよりも好ましいモデル選択性能

D-04 オンライン広告の実システムへの オンライン学習の適用

高木 潤, 田頭 幸浩, 野村 拓也, 岸本 忠士, 田島 玲 (ヤフー株式会社)

Yahoo! JAPAN の広告配信にオンライン学習を適用

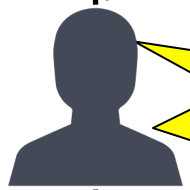
広告選択、配信

クリック率 予測

$$\text{CTR} = \frac{1}{1 + e^{-w_t^\top x_t}}$$



見る



クリック



w

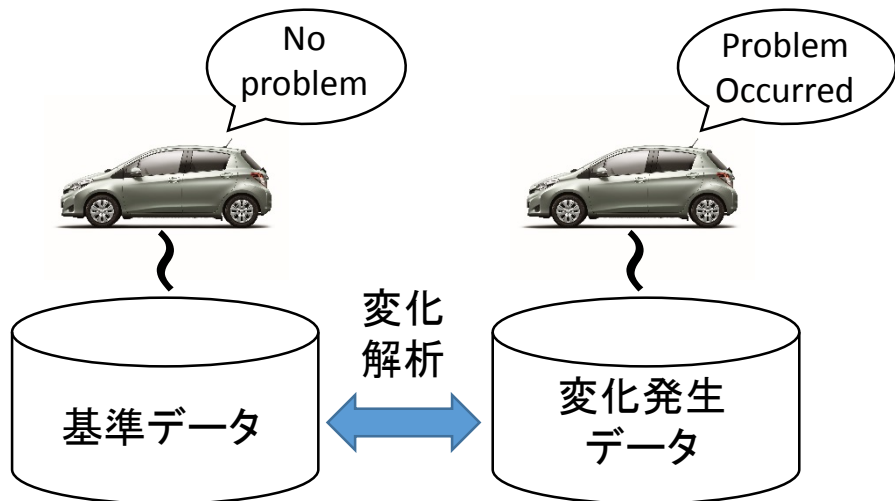
学習器

$$w_{t+1} = w_t - \eta \frac{\partial \mathcal{L}(w_t, x_t)}{\partial w_t}$$

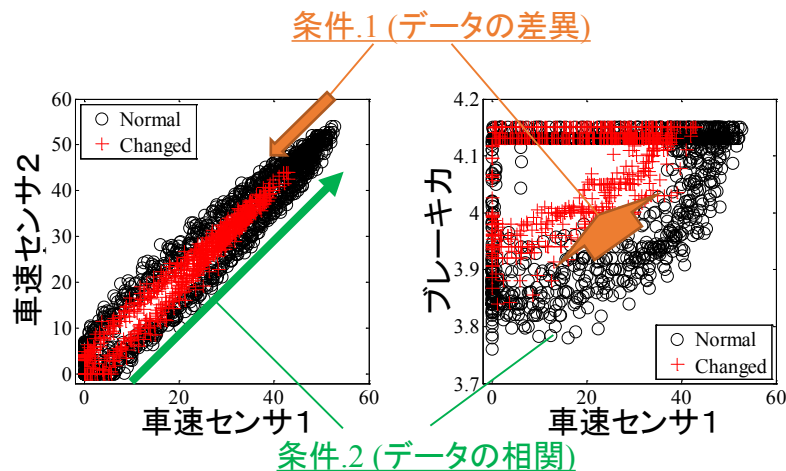
実システムに組み込む際の問題

- ・モデル選択
 - ・データがストリームで入ってくる
 - **cross-validation** できない
 - ・ onlineでvalidation??
 - ・ 適応的にハイパーパラメータ調整??
- ・確率値の予測がしたい
 - ・単純な2値分類手法は使えない
 - **confidence-weighted**等は不可

変化要因の可視化



散布図による要因の可視化



ハイパーパラメータ自動設定

$$\begin{cases} \Pr(\omega) = N(\omega \mid \mu = 0, \sigma = \omega_E) \\ \text{s.t. } \omega \neq 0 \end{cases}$$

常に高性能を維持するハイパーパラメータ
設定指針の導出

カーネル法とほぼ同等性能かつ高速な計算

特性カーネルに基づく手法との高い順位相関

データ	Ionosphere	Wdbc	Statlog	Optdigits
提案手法	0.9653	0.9786	0.9661	0.9511

計算時間の大幅削減

データ	Ionosphere	Wdbc	Statlog	Optdigits
提案手法	2.650(53)	3.435(32)	51.72(0.9)	149.5(1.0)
特性カーネル利用	4.996(100)	10.89(100)	5528(100)	14701(100)

パラメータ設定なく、高性能で、高速かつ、散布図で誰でも分かる変化要因可視化

[D-12] Riemannian preconditioning for tensor completion

Hiroyuki Kasai (UEC, Technische Universität München)

Bamdev Mishra (University of Cambridge, University of Liège,
Amazon Development Centre India)

- ▶ Propose a novel **Riemannian metric** that exploits both the **symmetry** present in Tucker decomposition and the **least-squares** structure of problem.
- ▶ The specific metric allows to use the versatile framework of **Riemannian optimization** on **quotient manifolds** to develop a **Riemannian preconditioned** nonlinear CG algorithm.
- ▶ Numerical comparisons suggest that our proposed algorithm robustly outperforms state-of-the-art algorithms.

D-14 正則化経験誤差最小化問題に対する確率的分散縮小双対平均化法

村田 智也, 鈴木 大慈 (東京工業大学)

(学生優秀プレゼンテーション賞対象)

正則項付き経験誤差最小化問題

$$\min_{x \in \mathbb{R}^d} \frac{1}{n} \sum_{i=1}^n f_i(x) + R(x)$$

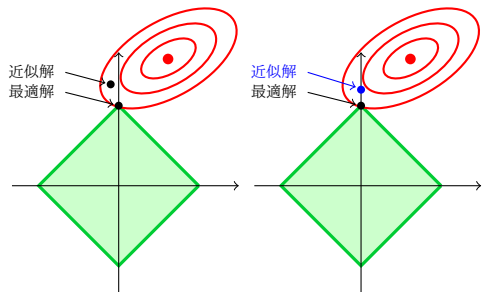
を解くための新しい確率的最適化アルゴリズムを提案。

利点

- R が非強凸のとき, オンライン型より速い収束レート.
- R が強凸のとき, 線形収束を達成.
- 数値実験で既存手法より高速な収束.
- 既存手法と比べ解がよりスパースに.

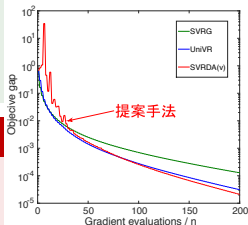
応用

Lasso, Elastic net 等の正則化に対して, より速く最適解のスパース構造を把握.

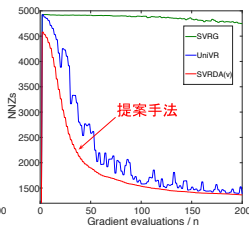


近似解がスパースでない

近似解がスパース



提案手法



提案手法

半順序つき非負値行列分解を用いた 交通リスクマイニング

D-16

李太斗 (東大), 松島 慎 (東大), 山西 健司 (東大, CREST)

研究目的

- 交通環境の効率的な改善
- 危険道路ランキングの作成

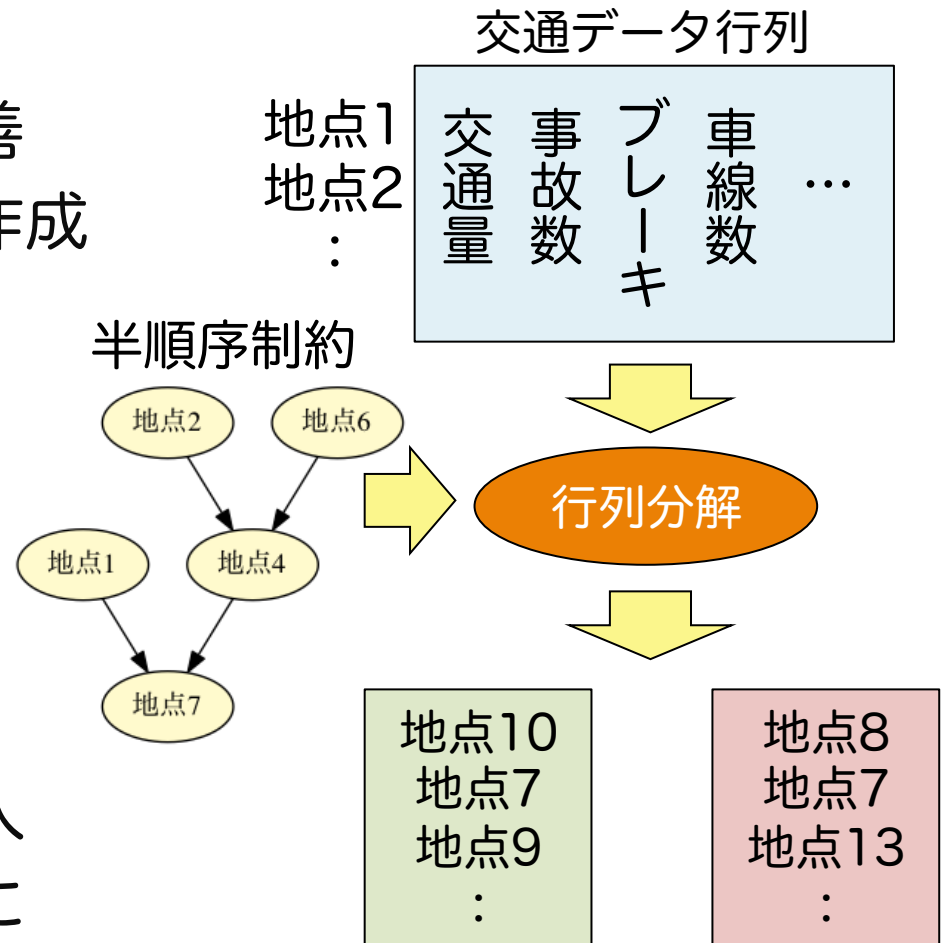
着眼点

- 非負値のデータから危険因子を抽出
→ **非負値行列分解**

提案手法

- データ間に**半順序**を導入
- 半順序関係を保つように行列分解

本研究はJST-CRESTの一部として行われた



複数のランキングを生成

学生優秀プレゼンテーション賞対象

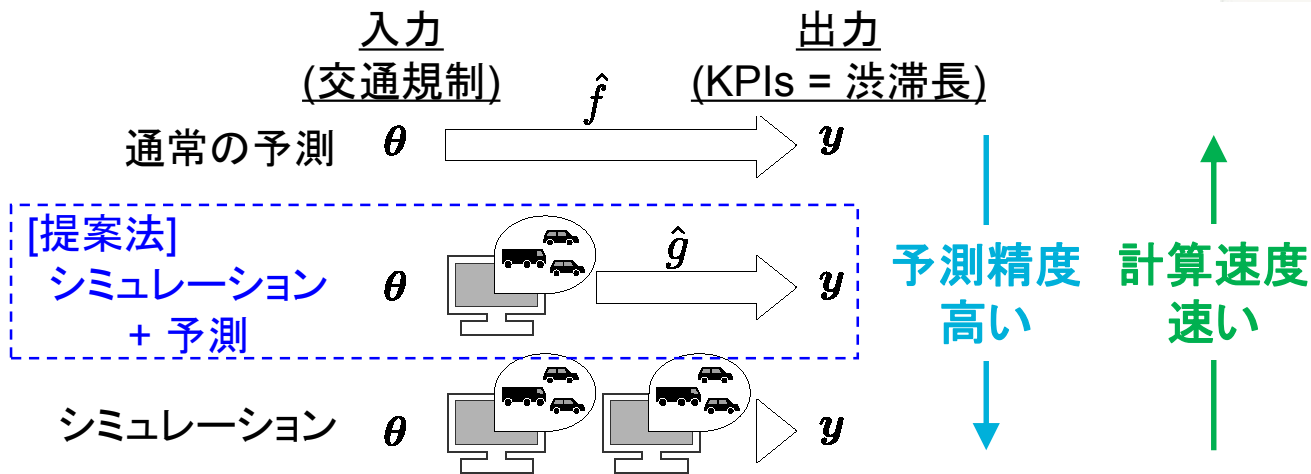
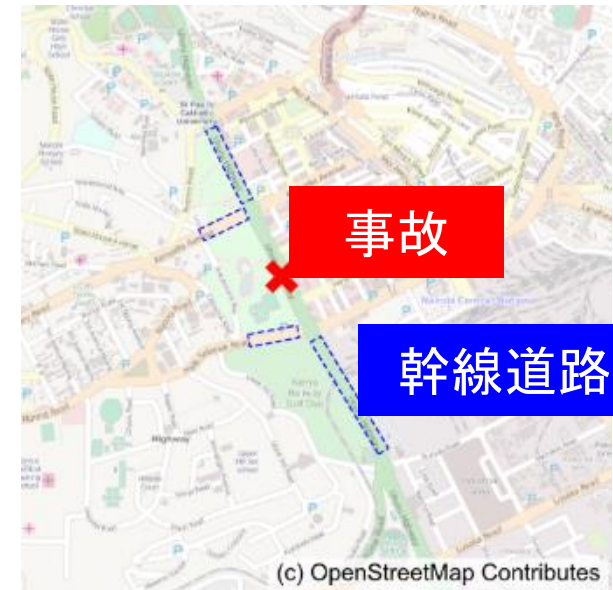
シミュレーションの計算過程からの結果KPI予測

D-18

原聡 (IBM東京基礎研究所)、森村哲郎 (IBM東京基礎研究所)

ルディー・レイモンド (JST/CREST)、牟田英正 (IBM Research - Australia)

- 問題例: 交通渋滞の予測 / 解決策の探索
 - 各幹線道路にどんな交通規制を敷いたら渋滞が減るか?
 - ・ 良い交通規制を効率よく探したい。
 - ・ 全パターンをシミュレーションするのは非現実的。
- 提案法: シミュレーション過程から結果KPIを予測
 - ✓ 予測精度向上: 予測とシミュレーションを組み合わせる。
 - 通常の予測よりも高精度なKPI計算が可能
 - ✓ 計算速度向上: 一部のシミュレーションを予測で代替する。
 - シミュレーション単体よりも短い時間で計算が可能

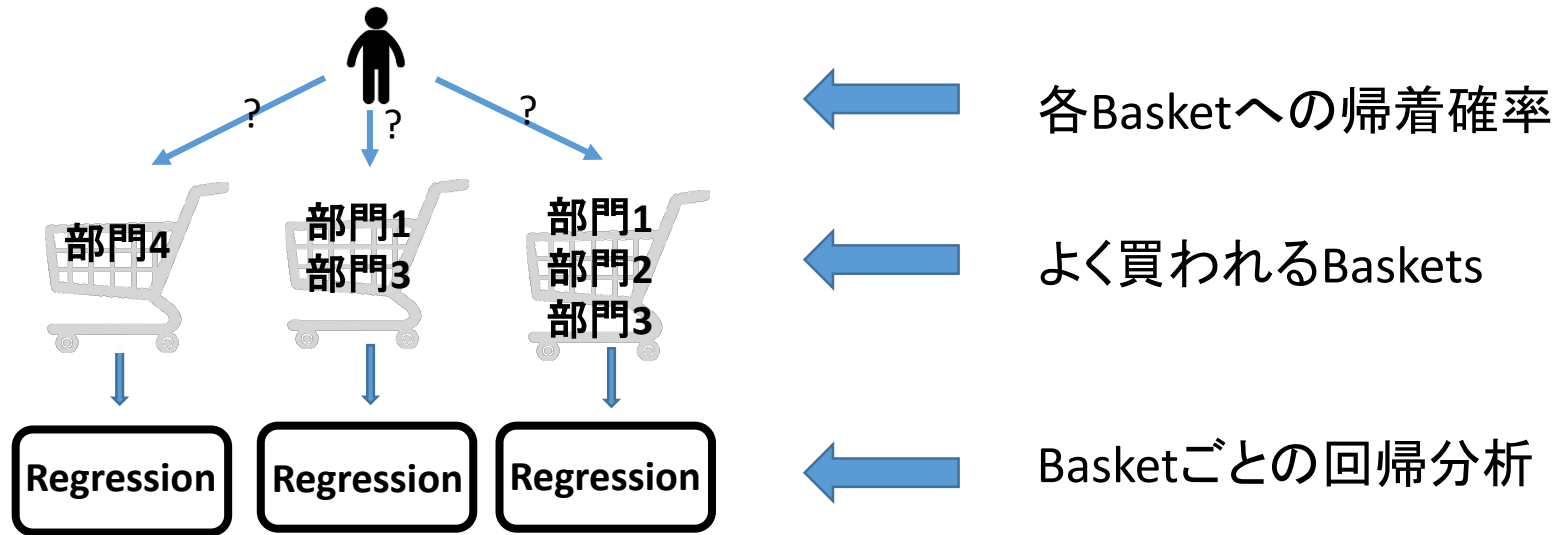


【提案法のポイント】
予測とシミュレーションのいいとこ取りをする。

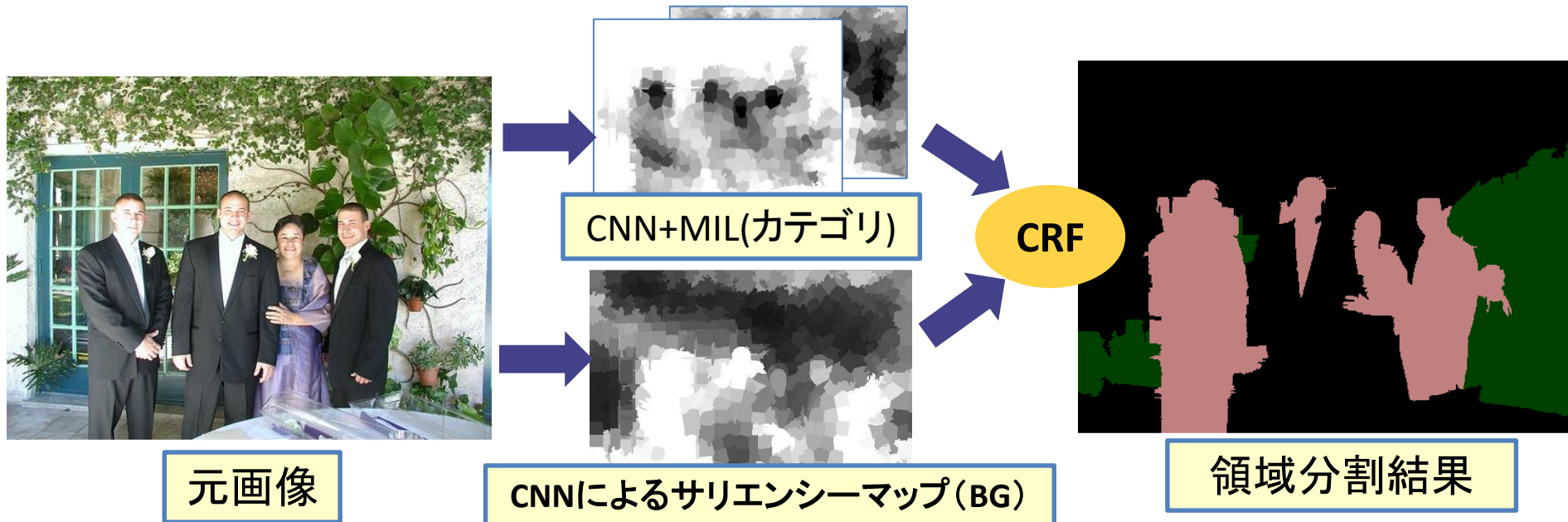
バスケット2パートモデルによる多部門医療支出データの分析

菅原慎矢(東大), ○吳天逸(東大, 修士1年), 山西健司(東大) **D-20**

目標	多部門の医療費への個人属性による影響の分析
既存研究	部門ごとの独立な分析
提案手法	バスケット分析と既存手法に基づいた多部門モデル
成果	多部門の同時購買関係の分析が可能 実データにおける知見(例:異なる購買パターンへの人種による影響)



D-22: CNNとMILを用いた 弱教師あり領域分割



学生優秀プレゼンテーション賞対象

電気通信大学大学院 総合情報学専攻

下田 和 柳井 啓司

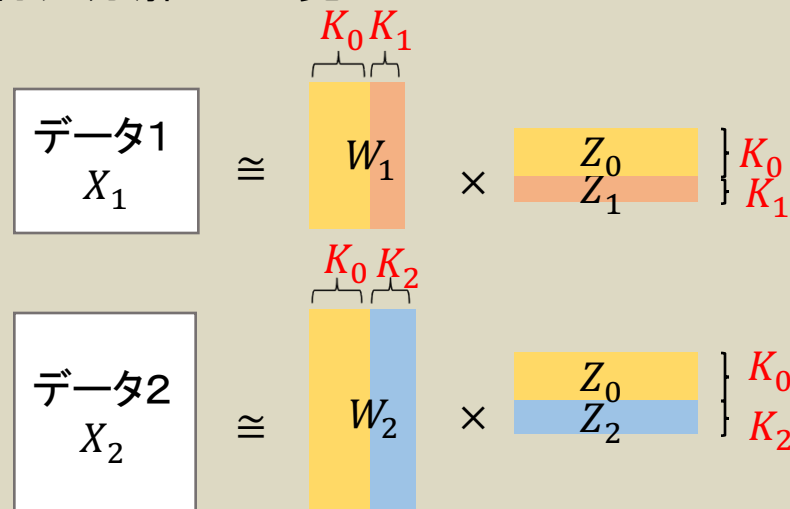
完全変数化NML符号長に基づく ベイズ正準相関分析のランク推定

中村 友彦(東大)、岩田 具治(NTT)、山西 健司(東大)

概要

- ベイズ正準相関分析(BCCA):
2つのデータ間で分解行列を共有する
マルチタスク学習
- 非共有要素を持つBCCA:
行列分解の**ランクが3つ必要で推定が困難**
- MDL規準に基づくランク推定手法を提案

行列分解として見たBCCA



提案手法

- 潜在変数モデルと解釈してモデルを正則化(完全変数化)
➡ NML符号長によるランク推定
- 関連度自動決定(ARD)を用いたノンパラメトリックベイズな既存手法よりも高いランク推定能力を確認



LEARNING VOLATILE TARGETS VIA STRATEGIC TRUNCATION

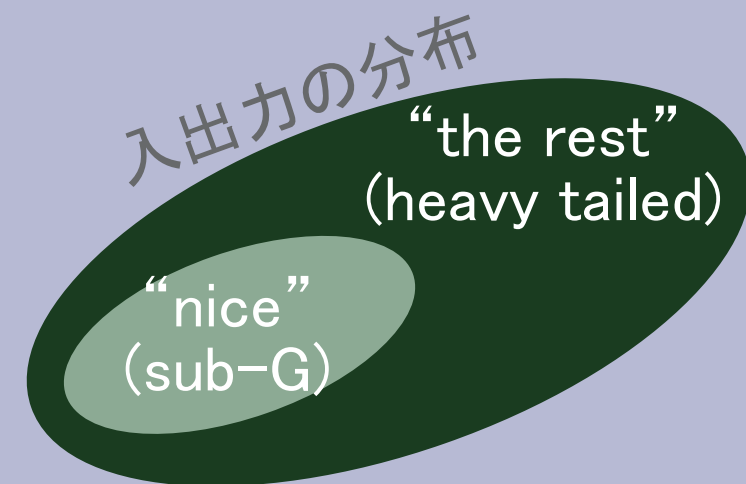
Matthew J. Holland, Kazushi Ikeda (NAIST)

D-28
学生優秀
プレゼン賞
対象

データの「暴れやすさ」に依る
大きな乖離 (ERM 系の**実践**・**解析**)

外れ値や非 Gauss
ノイズが潜伏 → 成績は顕著に
頑健性を欠く

ERM 系で推定
非 sub-G 許容 → 性能保障の枠組み、
事実上は崩壊...

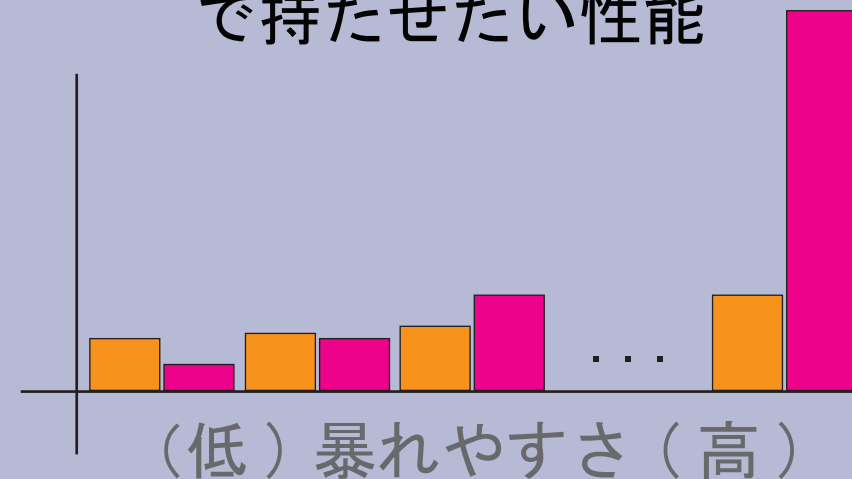


● → ●
目的関数「すり替え」
で持たせたい性能

本研究で知りたいこと

「**本命**」評価指標
 L を固定し、
heavy-tail を
許容しても...

... **目的関数**に着目し、
賢い「すり替え方」で
 L - 最良性能は
実現できる?



D-30: 同次ダイバージェンスとその応用

竹之内高志 (はこだて未来大学), 金森敬文 (名古屋大学)

モチベーション

離散確率モデルのパラメーター推定

⇒ 正規化項の計算が大変なことが多い

例: 隠れ変数を持つボルツマンマシン

$$\mathbf{x} = (\mathbf{x}_o, \mathbf{x}_h) \in \{+1, -1\}^{d+h}, \quad \bar{q}_{\boldsymbol{\theta}}(\mathbf{x}) = \frac{\sum_{\mathbf{x}_h} \exp(\boldsymbol{\theta}_1^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \boldsymbol{\theta}_2 \mathbf{x})}{Z_{\boldsymbol{\theta}}}$$

$$Z_{\boldsymbol{\theta}} = \sum_{\mathbf{x}} \exp(\boldsymbol{\theta}_1^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \boldsymbol{\theta}_2 \mathbf{x}): \text{正規化項}$$

提案法

(同次ダイバージェンス)+(経験分布による局所化)

⇒ 正規化項の計算を回避 & 漸近有効なパラメーター推定量

D-33

セーフスクリーニングを用いた 高次相互作用モデルによる分類アルゴリズム

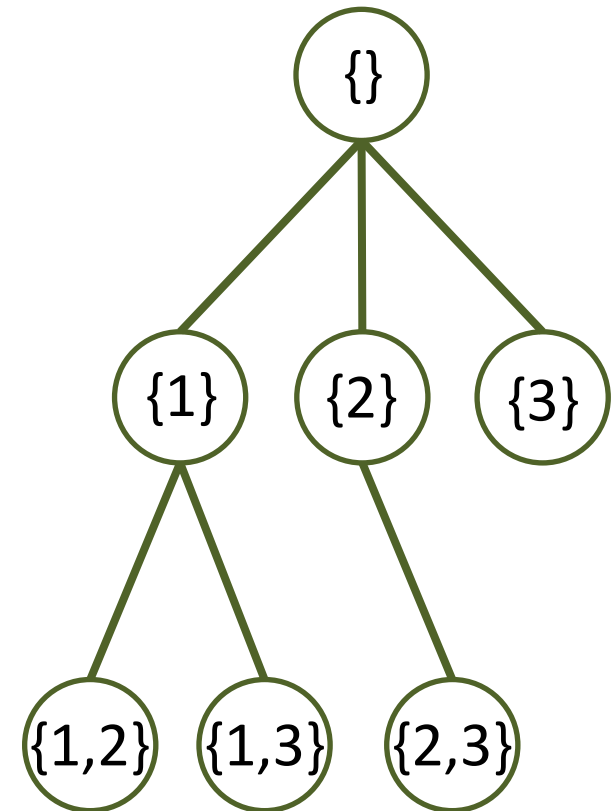
中川和也(名工大)・鈴木真矢(名工大)・烏山昌幸(名工大)・津田宏治(東大)・竹内一郎(名工大)

問題設定

- 高次相互作用モデルによる分類問題
- ナイーブな学習では全ての相互作用を取り扱うことが困難

提案法

- セーフスクリーニングにより学習に不要な相互作用を同定
- 木構造と枝刈りでセーフスクリーニングを効率化



多変量自己回帰モデルを利用した脳波時系列の欠測値推定

金子 拓光 (都立産技高専) 兼村 厚範 (産総研) 福永 修一 (都立産技高専)

生体計測では欠測が生じても継続して計測データを利用できるのが望ましい

同時刻を含む多変量自己回帰モデルにより欠測値推定する

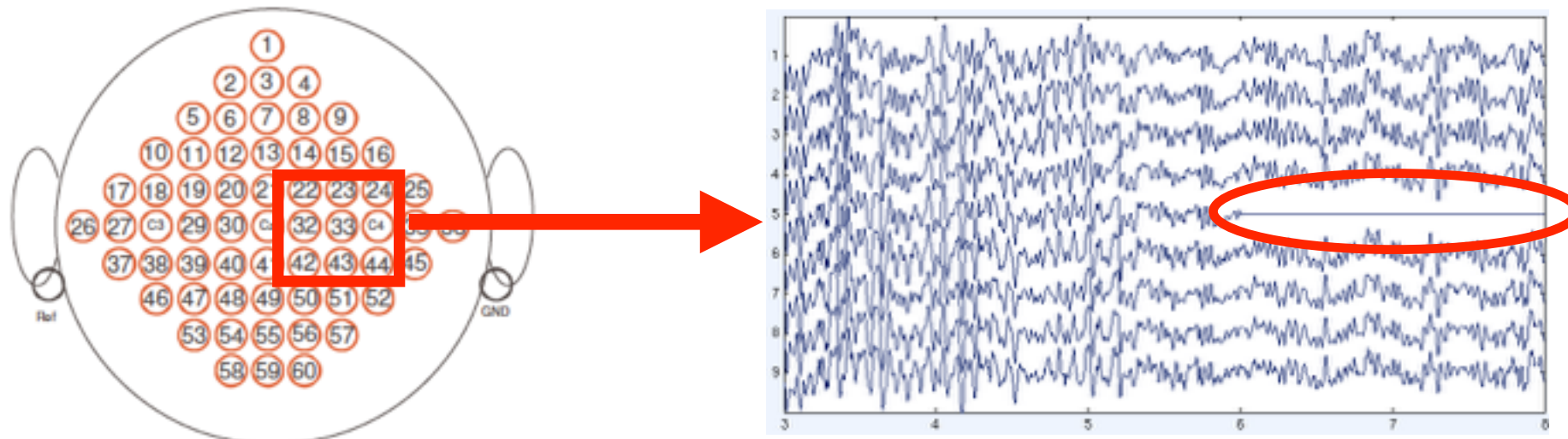


図1 電極配置図と欠測が生じた脳波時系列

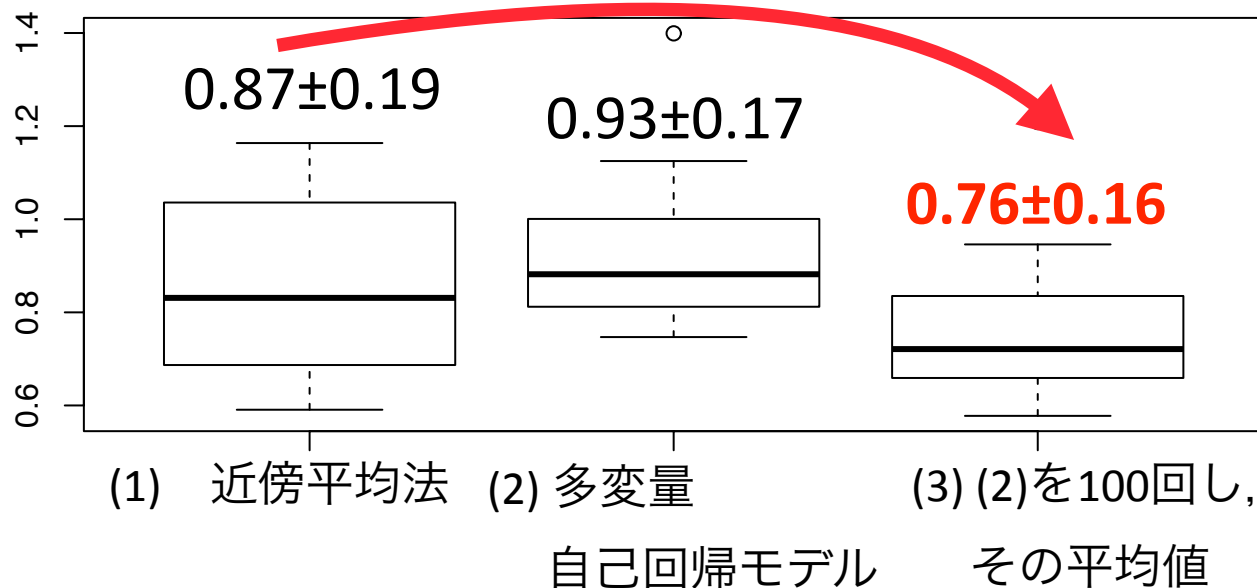


図2 各欠測値推定方法のRMSEの箱ヒゲ図

- 従来法よりも補完精度が向上
- 脳波の欠測値推定に有効
- 多重代入法へ発展可能

D37 Convergence Rates of Quasi Monte Carlo Integration under Misspecified Assumptions

金川元信(総研大), B. K. Sriperubudur (Penn. State), 福水健次(統数研)

タスク: 数値積分

$$\int_{[0,1]^d} f(x) dx \approx \frac{1}{n} \sum_{i=1}^n f(X_i)$$

解析的に計算できない

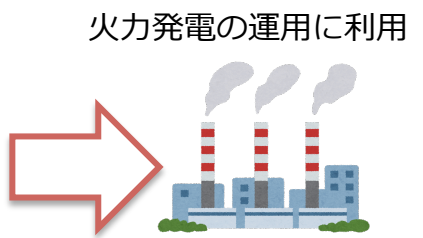
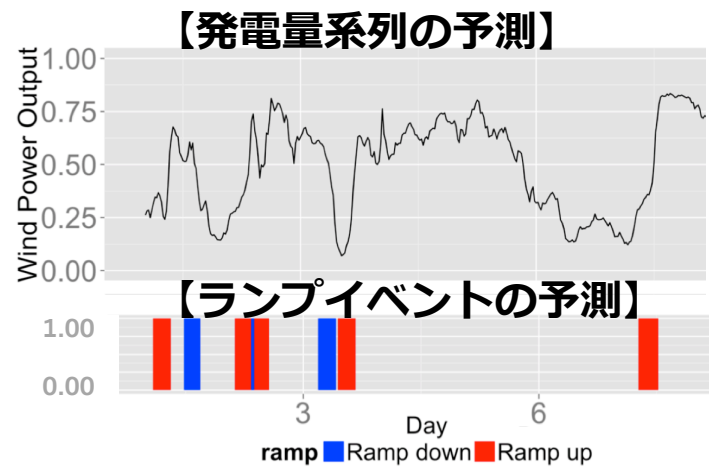
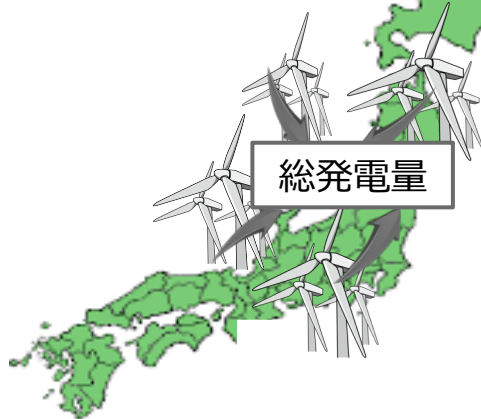
Quasi Monte Carlo (QMC)

X_1, \dots, X_n を**決定論的**に生成

➡ $O(n^{-1/2})$ よりも速い収束レート

本発表ではQMCに関する新たな理論結果を報告します

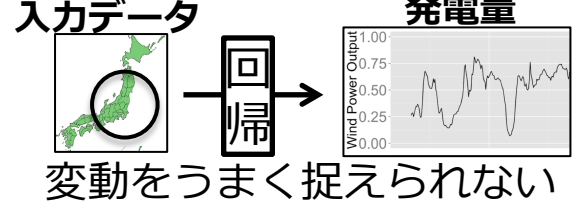
背景・目的 風力発電の急変（ランプ）を予測します。



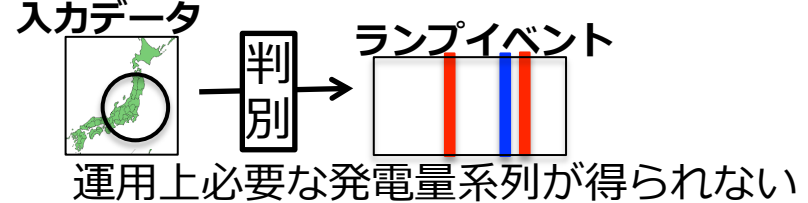
問題設定

• 発電量系列とランプイベントの両方を予測する。

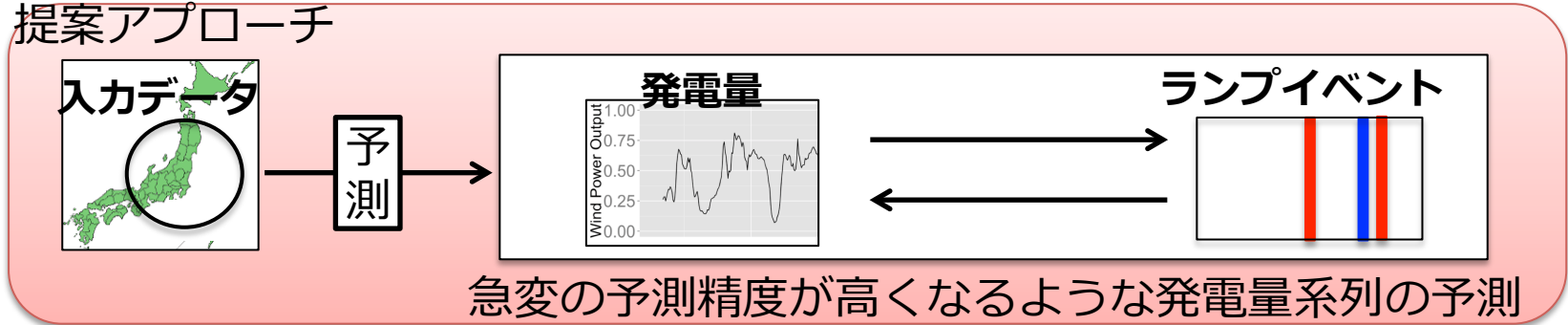
✓ 回帰アプローチ



✓ 判別アプローチ



✓ 提案アプローチ



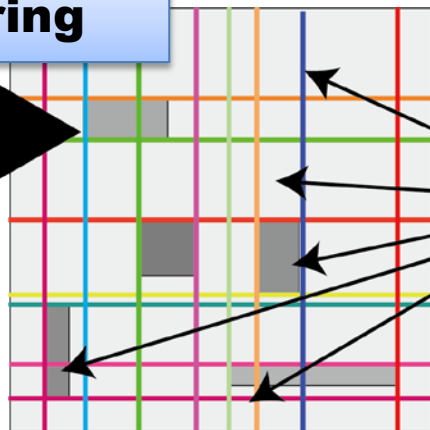
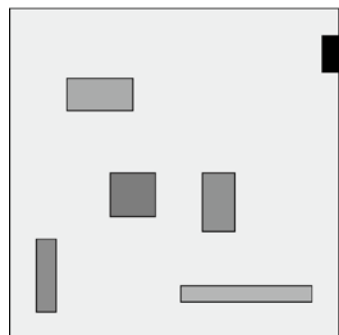
ランダムフォレストの近接行列を利用したアプローチを提案します。

D-41: Infinite bi-clusteringのための無限Plaidモデル



石黒勝彦 (NTT)、佐藤一誠 (東大)
中野允博、木村昭悟、上田修功 (NTT)

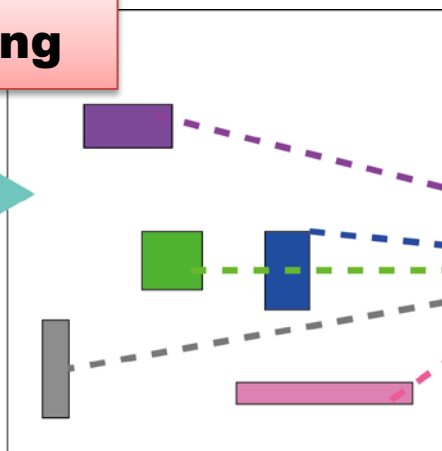
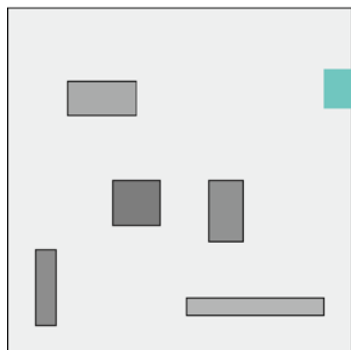
普通のbi-clustering



全部のブロックをいちいち目で確認しないといけない・・・
クラスタ数も決めないといけないし・・・



Infinite bi-clustering



解釈しやすい可能性が高いところだけ出てくると楽だな～
クラスタ数も自動的に決めてくれるし

完全変数化正規化最尤符号長に基づくCo-clustering

D-43

○尾亦範泰(東京大学), 山西健司(東京大学, CREST)

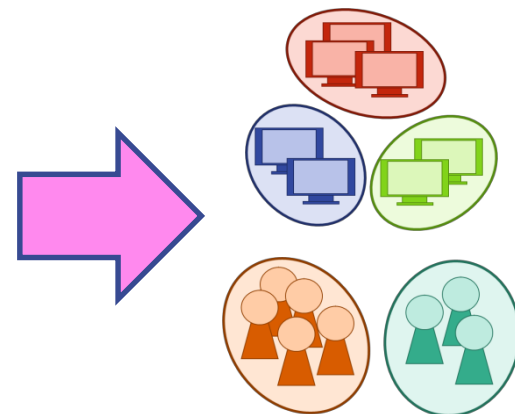
学生優秀プレゼンテーション賞対象

Co-clustering:

関係データを通じて異なる2つの集合をクラスタリング

課題: クラスタ数決定

	1	2	3	...	M
①	x	x	○	...	x
②	x	○	x	...	x
③	○	○	x	...	○
⋮	⋮	⋮	⋮	⋮	⋮
N	x	x	x	...	○



提案手法: 完全変数化正規化最尤符号長の適用

- 一般化EMアルゴリズムによる**最尤推定法の導出**
- 正規化項のRissanen漸近近似**計算式の導出**

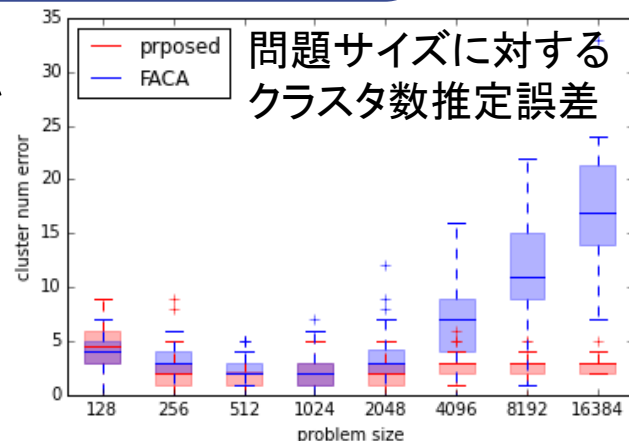
人工データ: 既存手法に対して有効性を確認

実データ: (提供: (株)博報堂)

広告業界の知識の再発見

(e.g. 主婦層に視聴率の高い番組群の発見)

本研究はJST-CRESTの一部として行われた

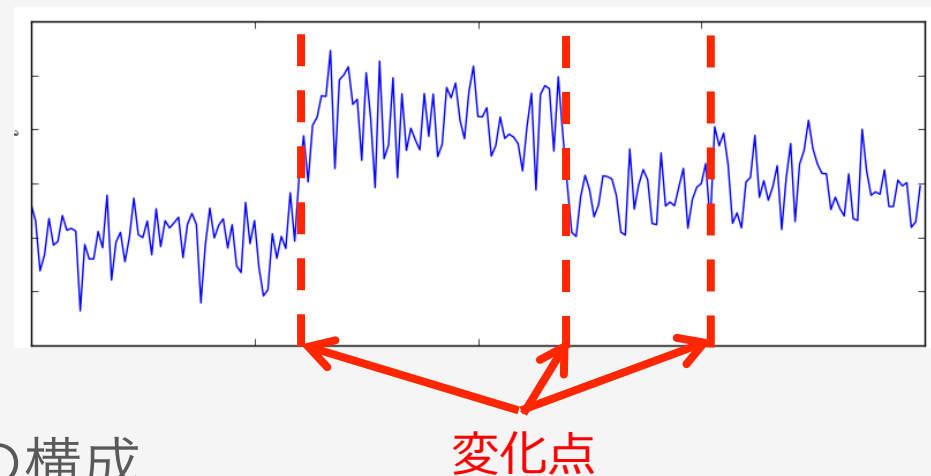


研究背景

- **変化点同定**：時系列の不定数個の変化位置を推定する問題
- 課題：理論保証と効率的な計算の両立

本研究

- 正規化最大尤度 (NML) を最大化
- 推定の**一致性**・**誤り確率**を評価
- 効率的な**オンラインアルゴリズム**の構成

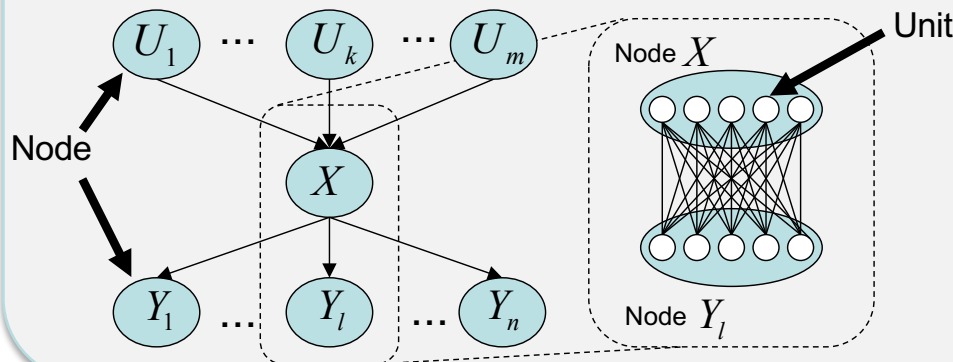


D-47: 大脳皮質モデルBESOMのクラスタ分散化と GPGPU並列化

中田秀基^{1,2}、黎明曦^{2,1}、井上辰彦^{1,3}、一杉裕志¹

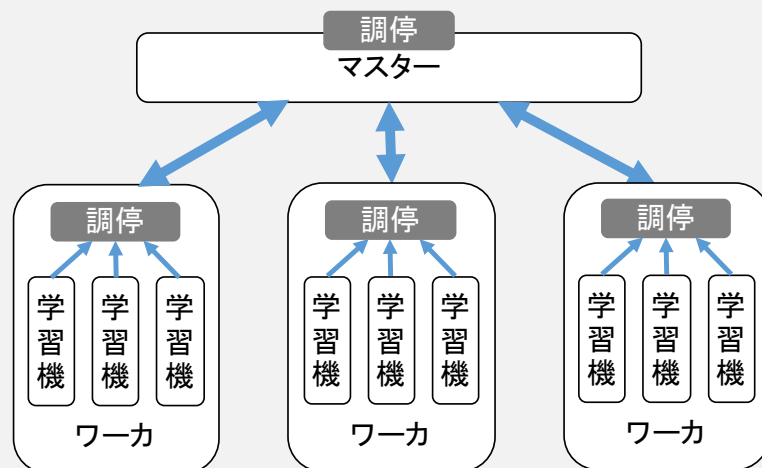
1 産業技術総合研究所人工知能研究センター、2 筑波大学、3 株式会社創夢

大脳皮質モデルBESOM

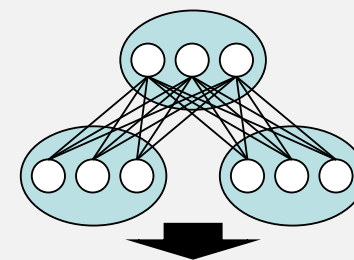


- ベイジアンネットワークによる、大脳皮質上のマクロコラム間ネットワークのモデル化
- ある仮定を導入することで、計算量をノード数のオーダーにおさえているが、ニューラルネットワークと比較すると計算量が多い
- 構造がノードとユニットの複層となっているため、計算が不均質

クラスタ分散化



GPGPU並列化



- データ構造をフラットに変更することによってGPGPUでの効率的な実装を可能に

D-49 教師生徒型強化学習のための複雑ネットワーク

石井良 鬘谷拓未 須鎗弘樹(千葉大学)

教師生徒型強化学習



教師が生徒にアドバイスしながら強化学習[1]

学習効率が大幅に上昇
生徒が教師の性能を上回ることもある

アドバイスのタイミング

- タイミングによって大きく学習効率が異なる
- 現在、5つの方法が提案されている

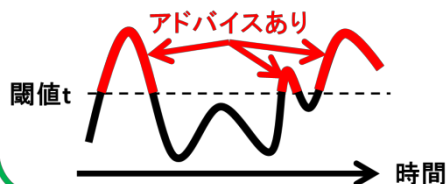
Early Advising

序盤にアドバイス



Important Advising

重要度が閾値t以上ならアドバイス

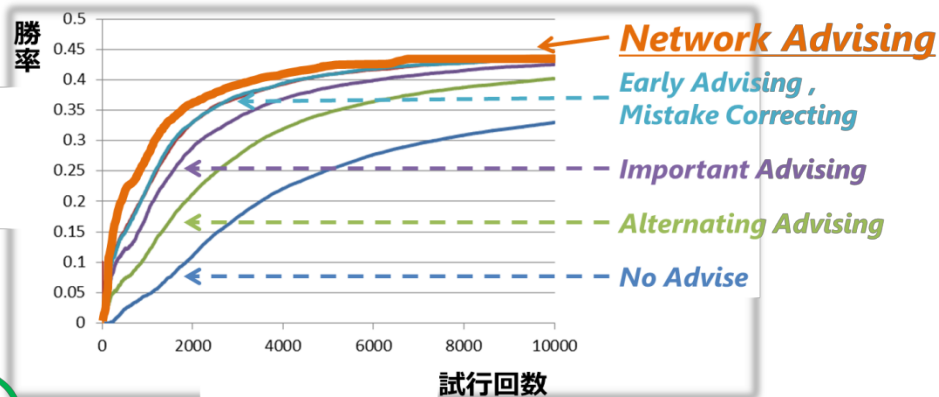


複雑ネットワークと教師生徒型強化学習

囲碁の行動を複雑ネットワークで表現[2]

強いネットワークと弱いネットワークに差異

複雑ネットワークの差異を利用して、新たなアドバイス、Network Advisingを提案



従来手法と比べ、最も高い学習効率

[1] M. E. Taylor, N. Carboni, A. Fachantidis, I. Vlahavas, L. Torrey, Reinforcement learning agents providing advice in complex video games, Connection Science, vol.26, pp45-63 (2014).

[2] B. Georgeot, O. Giraud, The game of go as a complex network, EPL, vol.97, 68002 (2012).

Gibbs事後分布の (ε, δ) -差分プライバシー

南賢太郎 (東大), 荒井ひろみ (東大), 佐藤一誠 (東大), 中川裕志 (東大)

ベイズ事後分布

$$p(\theta | x^n) = \frac{\prod_{i=1}^n p(x_i | \theta) \pi(\theta)}{\int \prod_{i=1}^n p(x_i | \theta) \pi(\theta) d\theta}$$

$$\Pi_i = \exp(\Sigma_i \log)$$

Gibbs事後分布

$\ell(\theta, x)$ 損失関数、 β スケールパラメータ

$$G_\beta(\theta | x^n) = \frac{\exp(-\beta \sum_{i=1}^n \ell(\theta, x_i)) \pi(\theta)}{\int \exp(-\beta \sum_{i=1}^n \ell(\theta, x_i)) \pi(\theta) d\theta}$$

本研究の問題

「Gibbs事後分布は (ε, δ) -差分プライバシーを満たすか？」

→ 主定理：(適当な条件のもと) **Yes**

キーワード：プライバシー保護、測度集中理論、PAC-Bayes

Expert advice による前立腺特異抗原の時系列予測

○森野佳生(東大), 平田祥人(東大),
林典宏(高岡市民病院), 颯川晋(慈恵医大), 合原一幸(東大)

研究の背景:

- ・ 前立腺特異抗原(PSA): 前立腺癌の進行度合いを精度良く測れるマーカー.
- ・ PSAの時系列の高精度予測: 臨床上非常に重要な問題.

先行研究:

- ・ PSA のダイナミクスを記述する様々な数理モデルが存在.
- ・ 最近, 我々は expert advice を拡張した手法を提案し, PSAの時系列を少数データから予測できることを示した(K. Morino *et al.*, *Scientific Reports* **5**, 8953, 2015).

今回の発表:

- ・ Expert advice によるPSAの時系列予測研究の最近の発展について報告する.

D-55

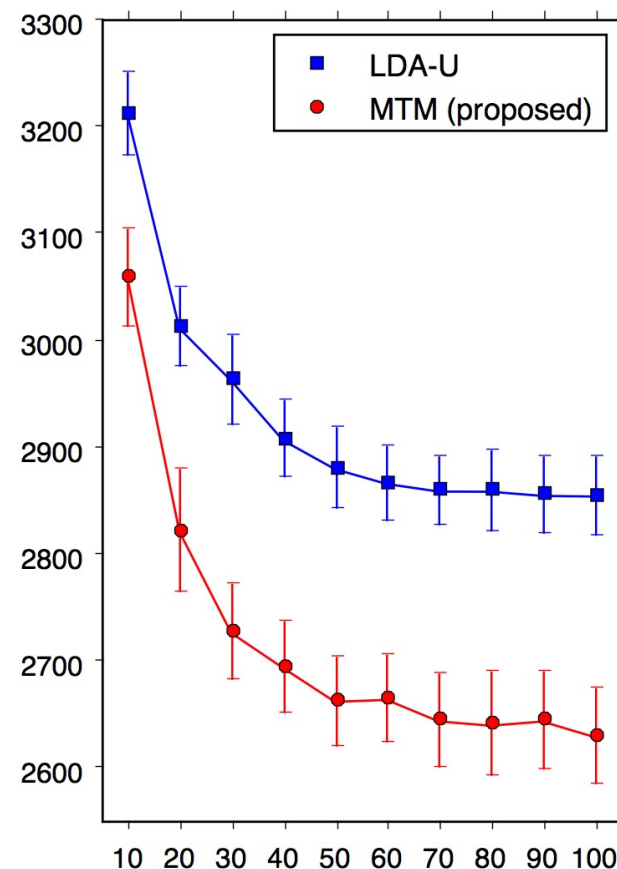
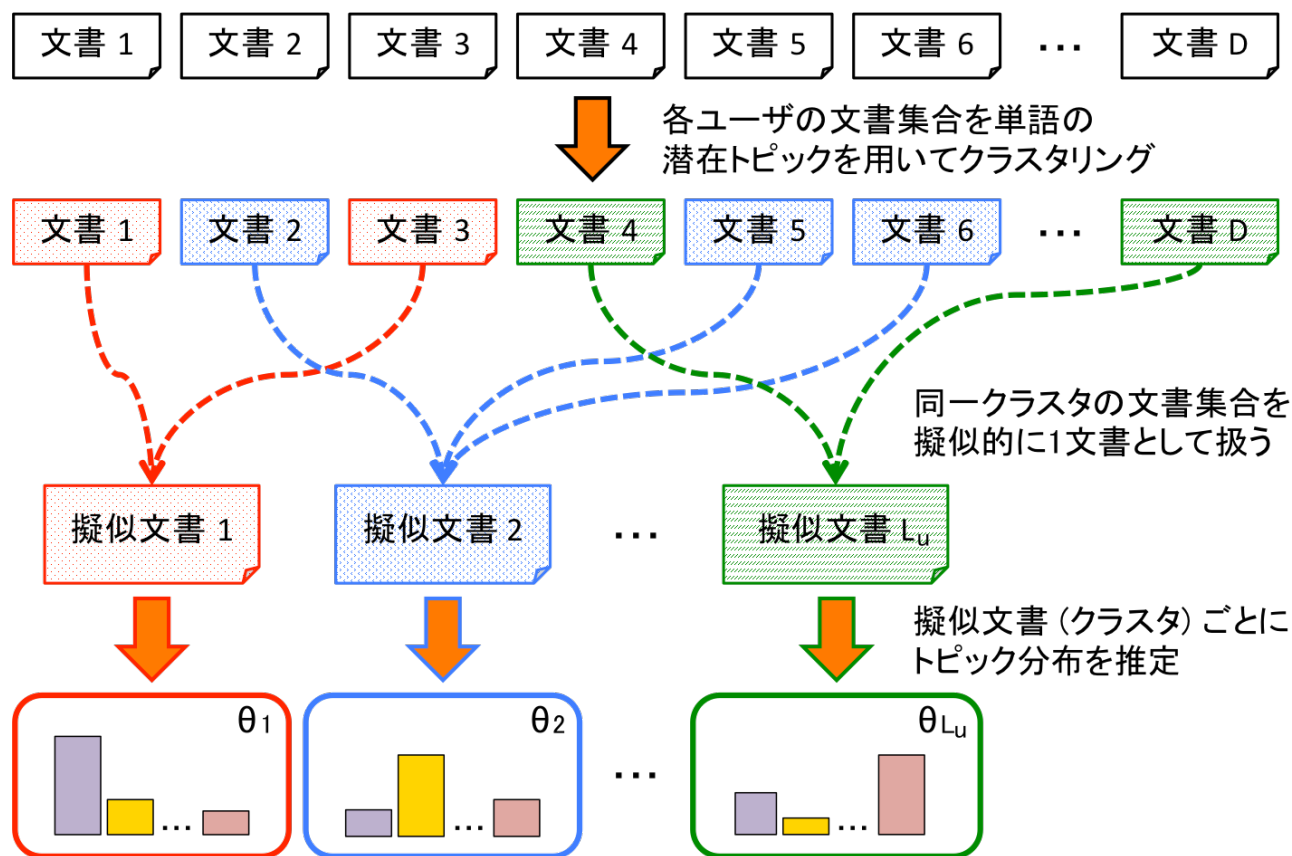
Topic Models for Analyzing Short Text Documents with User and Time Information

学生優秀プレゼンテーション賞対象

今井優作 (NAIST)

岩田具治, 澤田宏, 山田武士 (NTT)

各ユーザの文書の潜在クラスと単語の潜在トピックを推定



D-57: 医学・医療用疑似データセットの構築に向けて

城 真範・赤穂 昭太郎 / 産総研人間情報 2015/11/27@つくば

●問題意識:

プライバシー保護のために元データの開示が困難→研究結果の相互比較や追試が難しい

学術目的外の教育用資料やプレゼンテーションにデータ利用できない

●解決手段:

パラメトリックに作った疑似データを基準データとして使うことで研究比較プレゼンにも使える

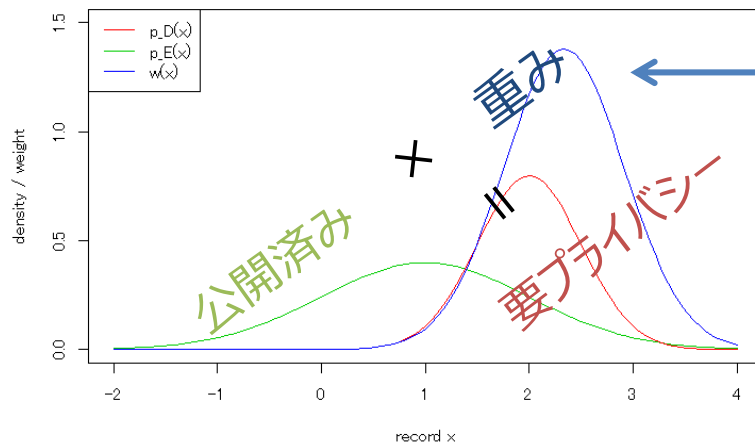
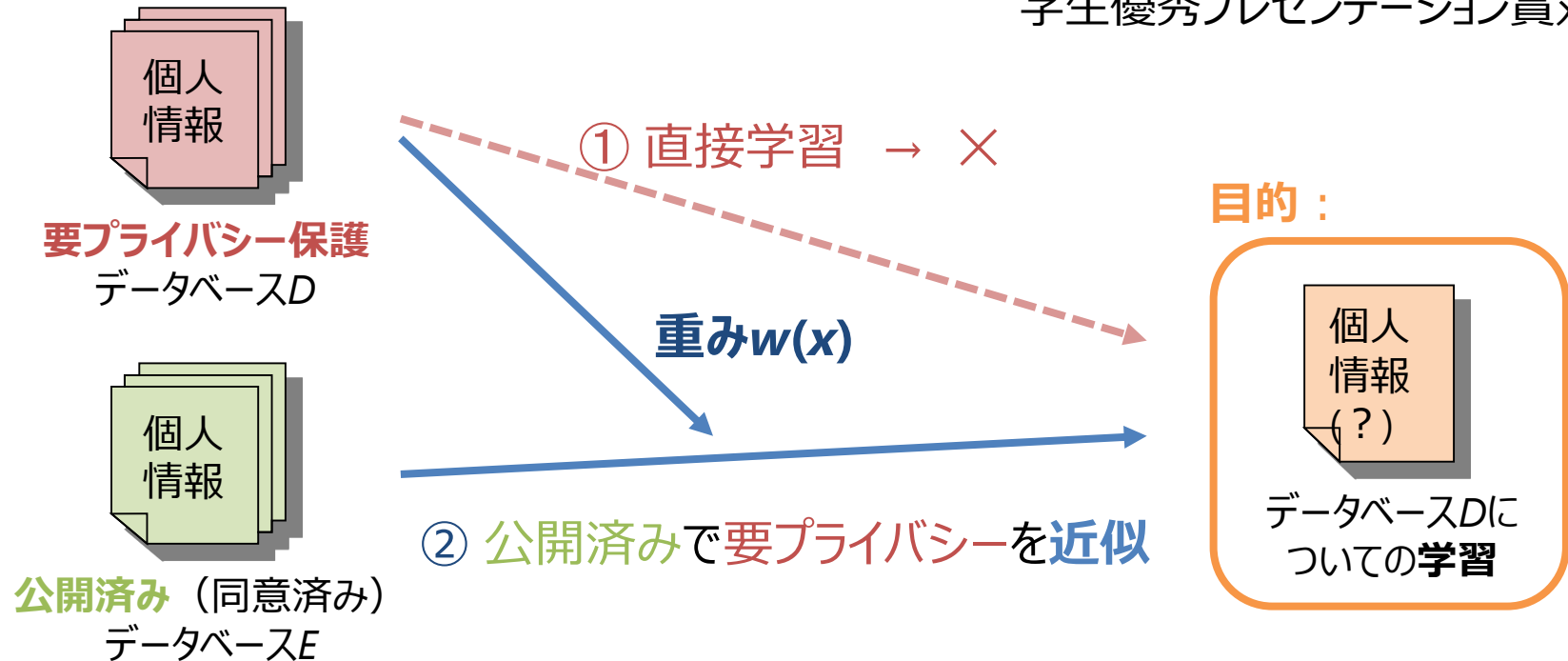
●実装:

数理的モデルが与えられている場合の疑似データ作成器 (Webアプリ) を作成

●課題:

実データの統計情報から血液検査値等、生体諸データのモデル構築を目指す
倫理審査申請中

学生優秀プレゼンテーション賞対象



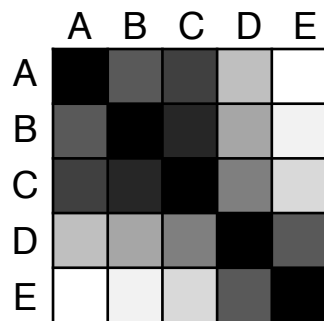
- ③ 重み関数 $w(x)$ を上手く推定
→ 最小二乗密度比推定 (uLSIF)
- ④ 重み関数自体も個人情報を含む
→ 差分プライバシーを満たすように
(プライバシー保証)

D-61 : Tree-likeな距離空間の特徴づけ

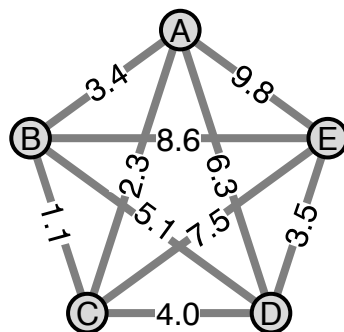
早水 桃子 (総合研究大学院大学) ・ 福水 健次 (統計数理研究所)

- Motivation : 与えられた距離情報を適切なグラフで表現したい。

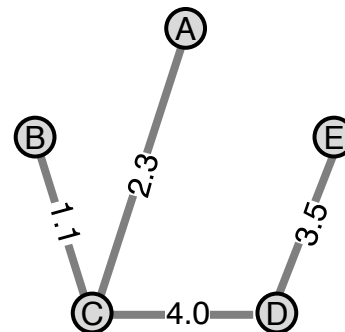
ペアワイズ距離



完全グラフ表現



辺数最少のグラフ表現



- Key questions

- 距離空間が木で表現可能であるための必要十分条件は？
- 距離空間を表現する木が存在するとき、その求め方は？
- 距離空間の「木らしさ」を評価するための尺度が得られないか？

エージェントモデルを用いた 緑内障病変の時空間伝播に関する予測手法

○友田恭輔(東大), 森野佳生(東大),
村田博史(東大), 朝岡亮(東大), 山西健司(東大, CREST)

学生優秀プレゼンテーション賞対象

目標と課題

- ・ 緑内障進行パターン^①の抽出をしたい
- ・ 緑内障の実データを扱う上では、観測間隔が異なる^②状況下で、視野上の各点間の遷移確率^③を推定する必要がある

概要

- ・ 時空間変化解析の分野では、時間的に等間隔な観測、場所的に近傍への遷移を前提とした手法が提案されてきた
- ・ 本研究では異なる時間間隔での観測から、近傍以外への遷移も捉える手法を提案する

提案手法

- ・ 遷移パターン^④を捉えるために隠れ変数を導入
- ・ エージェントが病気を運んでいくモデル^⑤を利用
- ・ 視野上の伝播のしやすさを同時に推定

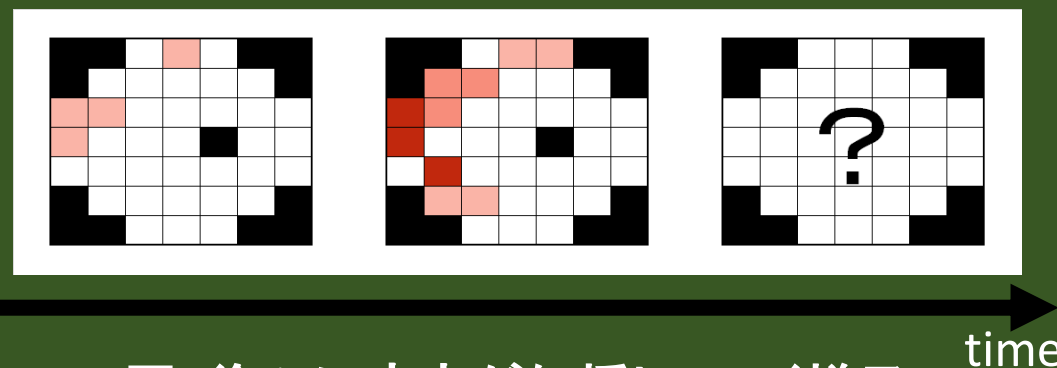
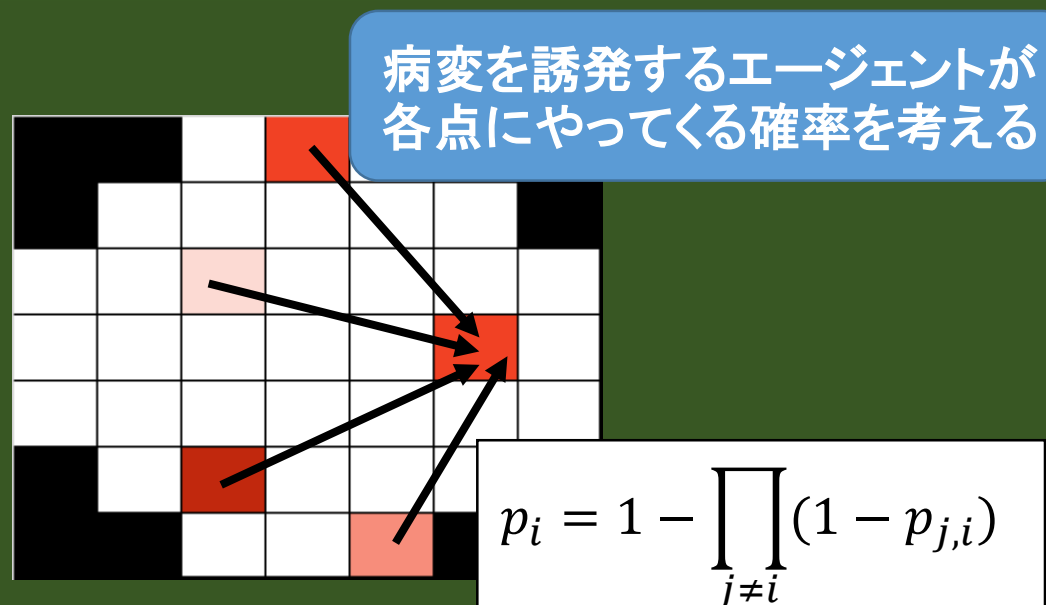


図: 徐々に病変が伝播していく様子



本研究はJST-CRESTの一部として行われた

相互情報量の推定における Minimax最適推定量

福地 一斗(筑波大), 佐久間 淳(筑波大/JST CREST)

独立性指標(相互情報量)の推定問題

未知分布のサンプルを元に未知分布に関する独立性指標(相互情報量)を推定する問題

サンプル分布

サンプル

独立性指標

\mathbb{P}

\longrightarrow

$(X_1, Y_1), \dots, (X_n, Y_n)$

\longrightarrow

$I_n \simeq I(X; Y)$

未知

既知


サンプルから推定

アプリケーション: 特徴選択, クラスタリング, 公平配慮学習

既存の結果

- 一定条件下で損失の上限が $O(\sqrt{1/n})$
- Minimax損失の下限は示されていない

貢献

- X, Y が有限の時のMinimax損失の下限が $\Omega(1/n)$
- 一定条件下で損失の上限が $O(1/n)$ の推定量の提案 

重み付きグラフに対するマイクロクラスタリングの 拡張とオークションデータへの応用

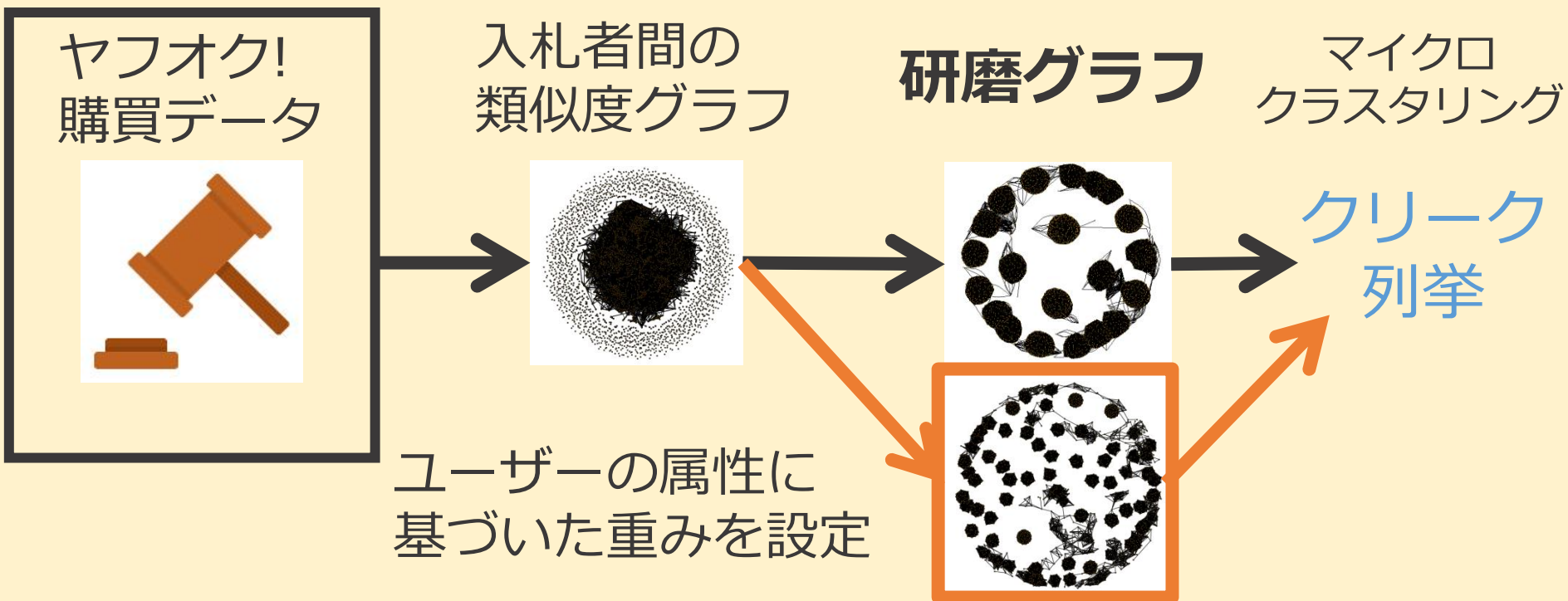
山崎 朋哉(京都大学), 清水 伸幸, 小林 隼人, 山内 智(ヤフー株式会社)

マイクロクラスタリング [T.Uno+ 15] →

- 重みなしグラフに対する高速なグラフクラスタリング手法

提案手法 →

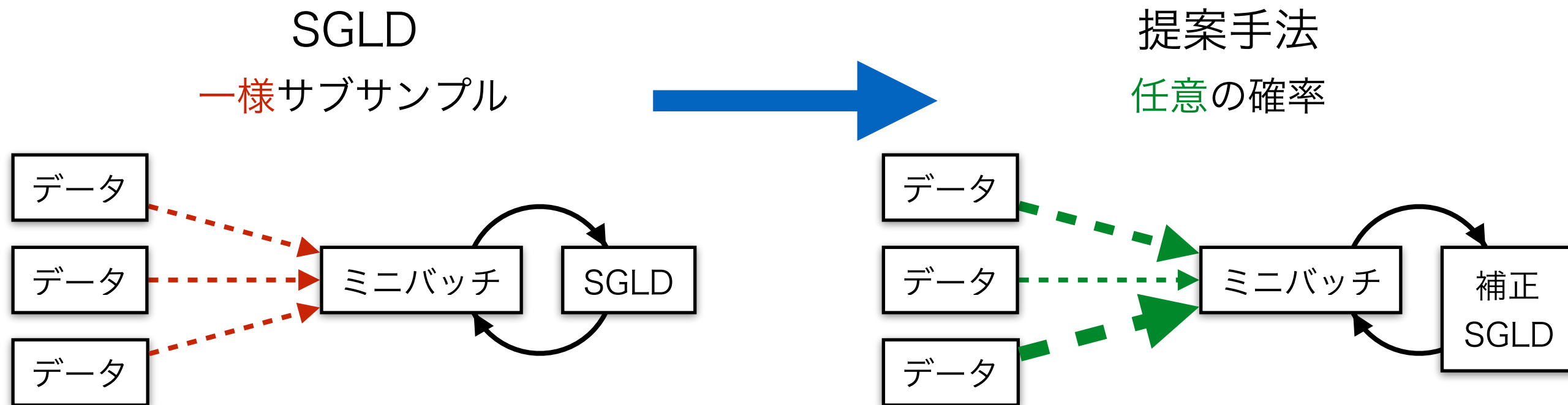
- マイクロクラスタリングを重み付きグラフへ適用可能に



D-69 非一様サブサンプルによるSGLD

横井創磨、佐藤一誠、中川裕志（東大）

学生優秀プレゼンテーション賞対象



利点1. 一様サブサンプルできない状況でも実行可能

確率質量の**99%**がデータの**1%**に集中 → 正確な事後分布

利点2. わざとサブサンプルを偏らせることで効率化

各データのコストに応じて選び方を変える → **2~5倍**の高速化

D-71

非線形テンソル学習手法の高速化と Yahoo!ショッピング購買金額予測への適用

金川平志郎¹・清水伸幸²・小林隼人²・田頭幸浩²・鈴木大慈¹

¹東京工業大学 ²ヤフー株式会社

学生優秀プレゼンテーション賞対象

非線形テンソルモデル
(変数の組の非線形な関係を記述)

$$y = \sum_{r=1}^{d^*} \prod_{k=1}^K f_r^{(k)}(x^{(k)}) + \epsilon$$

Gaussian process
priorで推定

ベイズ的アプローチ

→計算: Gibbs Sampling

・問題 = 計算時間

- ・サンプリングでの逆行列計算

・本発表

1. 高速化

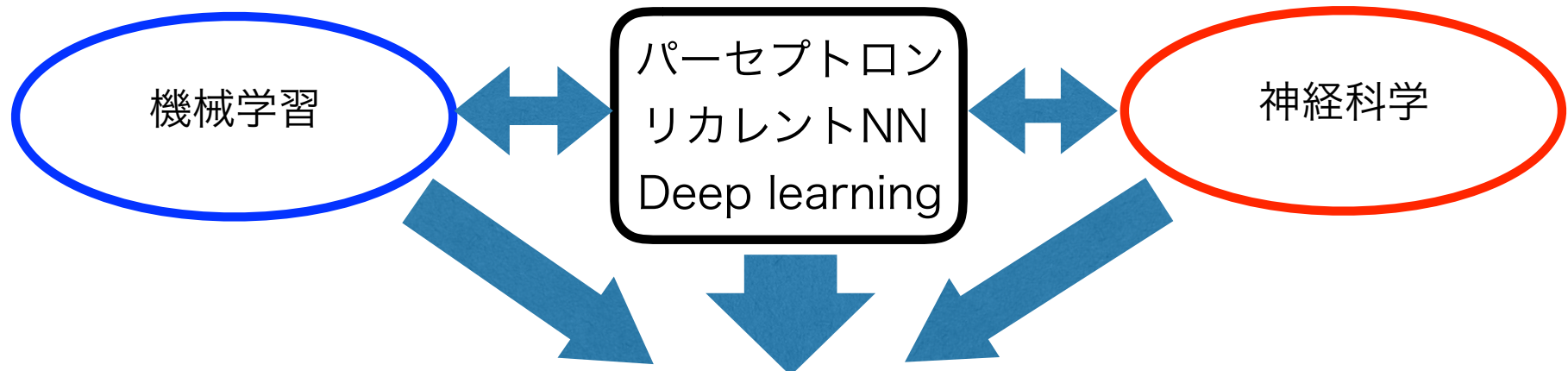
- ・Nyström近似による高速化
- ・高速化の課題

2. 実データへの適用

- ・Yahoo!ショッピングデータへの適用
- ・ユーザの特徴の非線形性

大規模ベイジアンネットワークによる ディープラーニング実用化に向けた課題の検討

佐野崇、川田正晃、一杉裕志 (産総研)



- ・ 大脳新皮質=統計的情報処理, e.g. 感覚統合, 運動制御, 言語, etc...
- ・ Belief propagation と類似する双方向結合
- ・ 大規模な隠れ変数を持つベイジアンネットワークとして大脳をモデル化→スケーラビリティの秘訣を明かす
- ・ 脳に学ぶことで、次の世代のdeep learningを目指す

D-75

混合ガウス過程を用いた時空間変数予測

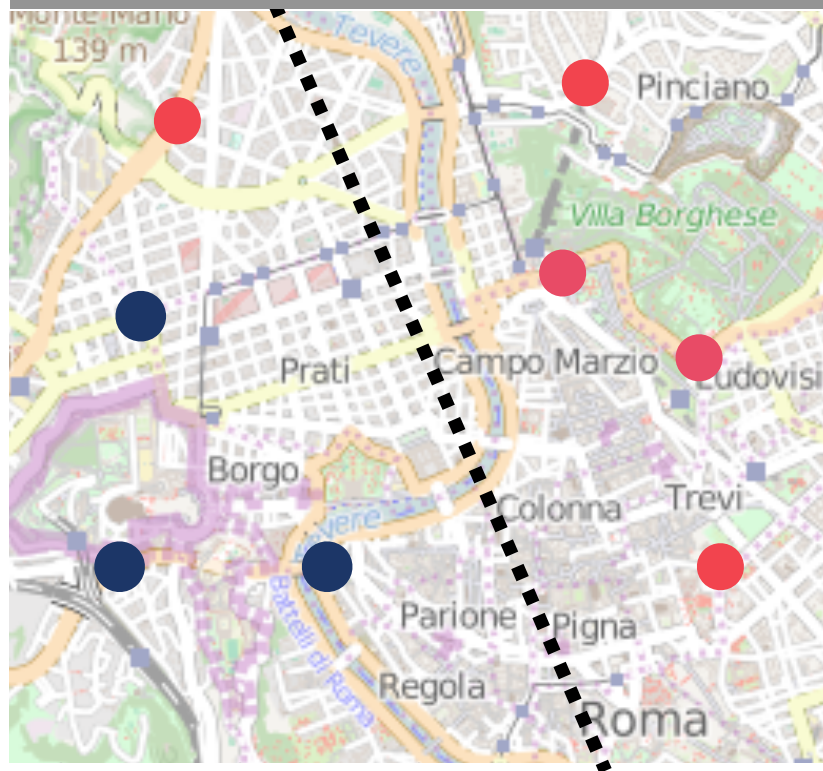
大川 真耶¹ 澤田 宏^{1,2} 上田 修功²

1. NTTサービスエボリューション研究所

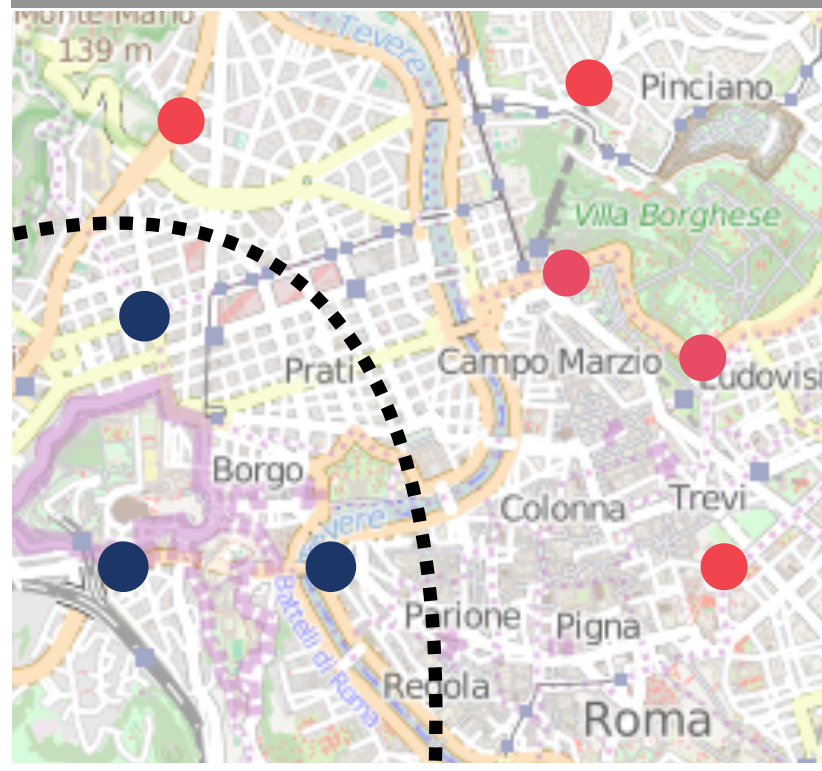
2. NTTコミュニケーション科学基礎研究所

- ◆ 混合ガウス過程を拡張したモデルを提案
- ◆ 提案手法は非線形なクラスタ抽出が可能

従来手法 (混合ガウス過程)



提案手法



Stiefel多様体における連続値入力RBMの最尤推定

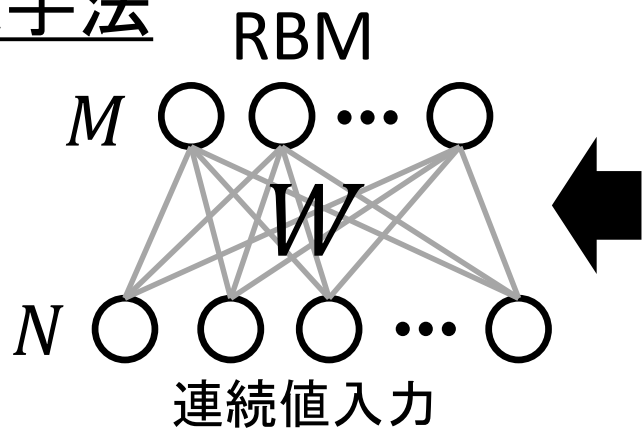
唐木田 亮 (東大), 岡田 真人 (東大), 甘利 俊一 (理研)

背景 RBMの最尤推定は解析計算が困難

尤度: MCMC (計算時間が長い)

勾配: Contrastive Divergence (収束の保証なし)

提案手法



直交制約

$$M \left[\underbrace{W}_{N} \right] W^T = I_M$$

測地線に沿った自然勾配でWを学習

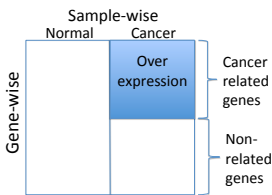
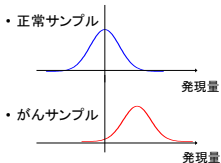
尤度と最尤勾配の解析計算が可能

尤度: 高速に計算 勾配: 最尤推定値に収束

A classification method based on nested normal mixture models with cancer outlier profile analysis

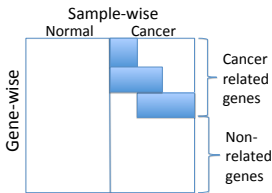
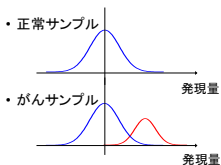
K. Matsui (名大/JST CREST), T. Oura (Eli Lilly Japan), S. Matsui (名大/JST CREST)

従来想定されてきた
がん関連遺伝子のプロファイル(分布)



- ▶ Nested mixture model によって outlier 構造を特定.
- ▶ サンプル (患者) の事後確率を導出.
- ▶ 確率的にがん/健常を判別.

Cancer outlier 型プロファイル

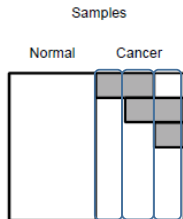


遺伝子発現量:

$x =$

x_1
 \vdots
 x_p

Genes



事後確率: $P_1(x)$ $P_2(x)$ $P_3(x)$

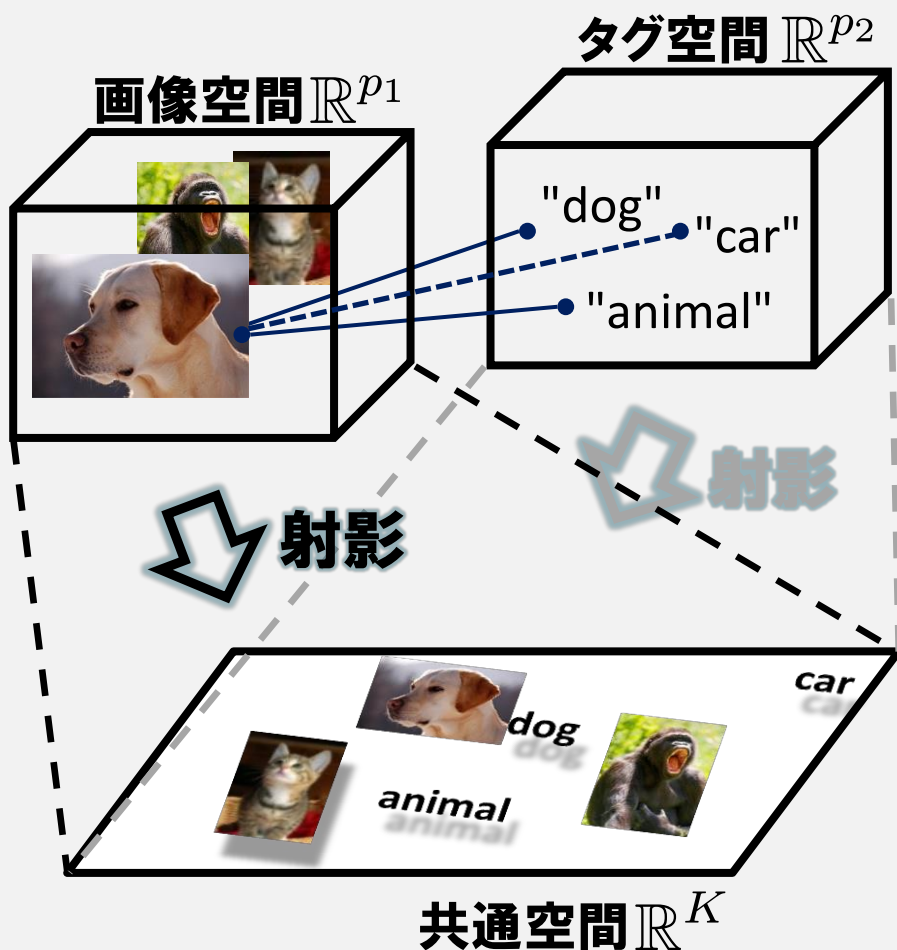
D-79 (学生優秀プレゼンテーション賞対象)

ロバストな多ドメインマッチング相関分析の提案と 画像認識への応用

大阪大学大学院基礎工学研究科

奥野 彰文

下平 英寿



X^{Img} , X^{Tag} : データ行列

W : タグ付けの情報
(マルチタグ可)

問題:

誤まったタグ付けが混入

目的:

射影行列をロバスト推定

\hat{A}^{Img} , \hat{A}^{Tag}

画像-タグの相互検索

D-81 逆強化学習の枠組みを用いた未知状況下の移動予測に関する検討

幸島 匡宏, 松林 達史, 澤田 宏(NTT)

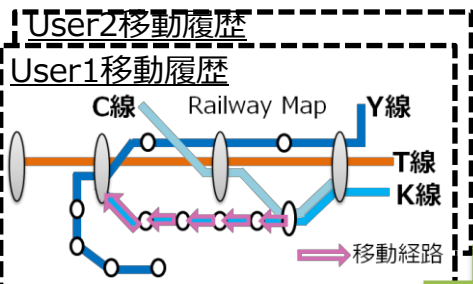


□ 自然災害・事故発生時などデータ収集時とは異なる未知の状況下における予測を行うという問題に取り組みます。

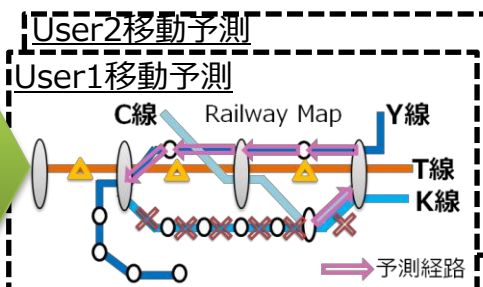
□ 逆強化学習と呼ばれる, ユーザの移動を説明する“報酬”をデータから推定する手法を用いたアプローチを発表します。

未知状況下予測の例: 事故時の移動予測

入力1: 事故未発生時のユーザ移動履歴



出力: 事故発生時のユーザ移動予測

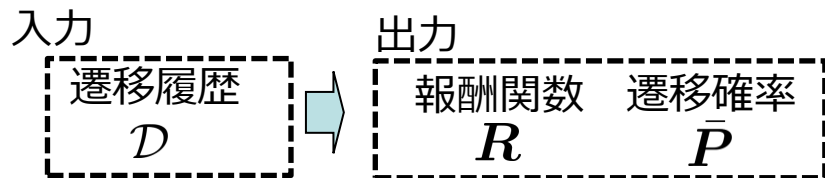


入力2: 事故情報

A鉄道会社の運行状況	
K線 上下線 全線	✕ 運転見合わせ 原因: 架線断線
T線 上下線 全線	△ 遅延 原因: 架線断線

逆強化学習を用いたアプローチ

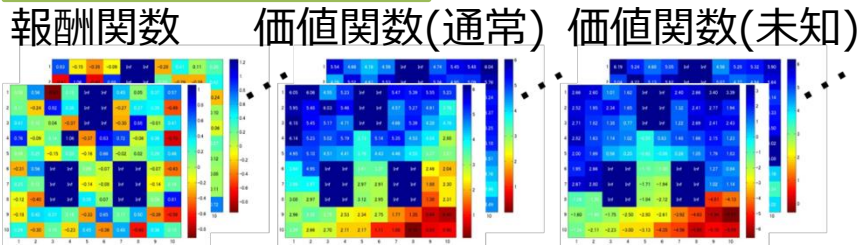
Step1: 報酬推定



Step2: 事故情報統合

Step3: 未知状況下遷移確率算出

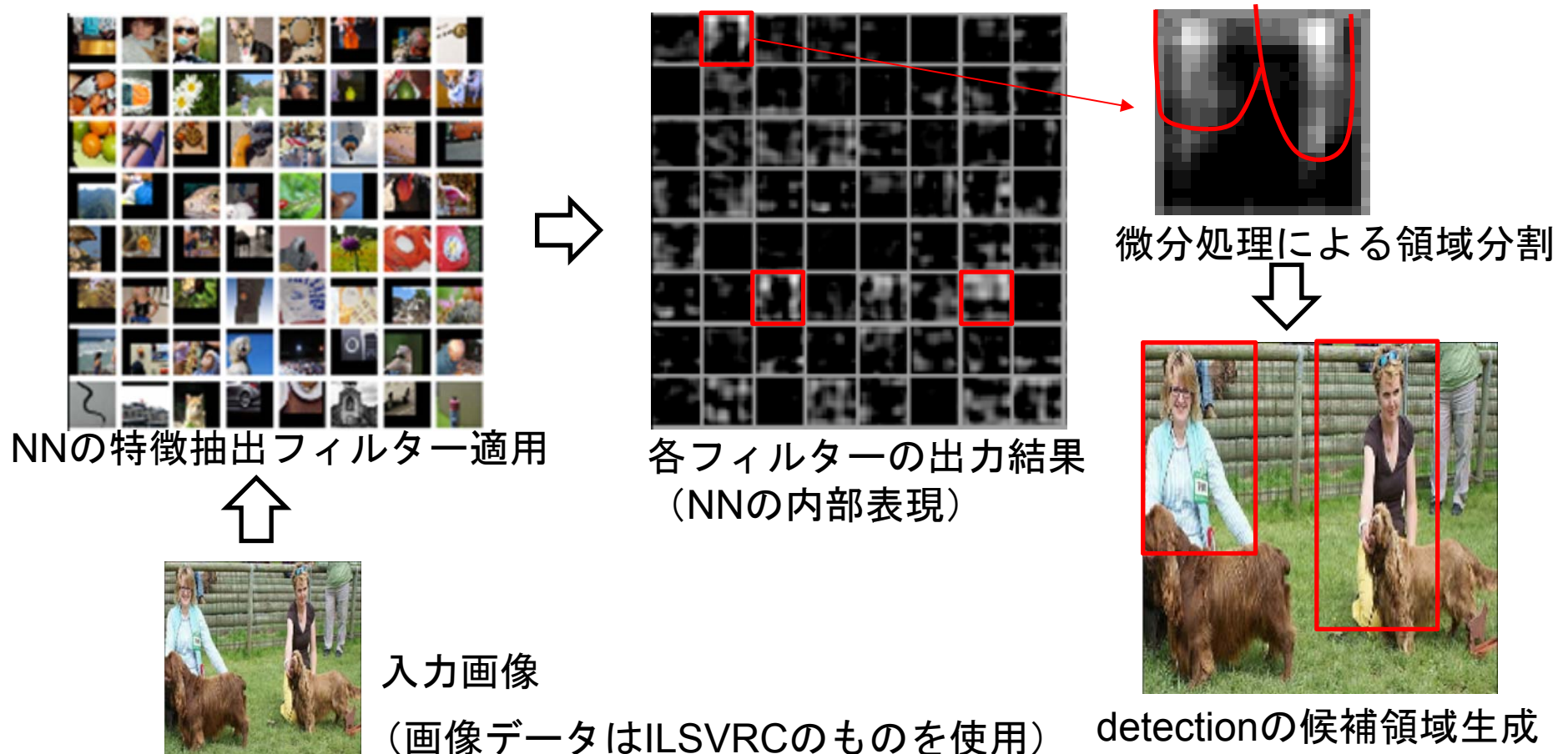
人工データ適用結果



D-82: ディープラーニングによる detectionの一手法

発表者：足田聡（株式会社リコー）

- ニューラルネットワークによって抽出された内部表現を利用してオブジェクトの候補領域を生成する手法を紹介



D-83 Latent Synonym Topic Models

Weihua Hu(UTokyo), Motoki Sato(NAIST), Hideki Mima(UTokyo), Jun'ichi Tsujii(AIST)

	LDA	Latent synonym Topic Models
Topic as a distribution over	words.	latent synonyms.
Topic inference depends on	co-occurrence of words.	co-occurrence of latent synonyms.
Vocabulary size is	large and sparse. (Usually 5000 - 10000)	small and dense. (Usually 500 - 1000)

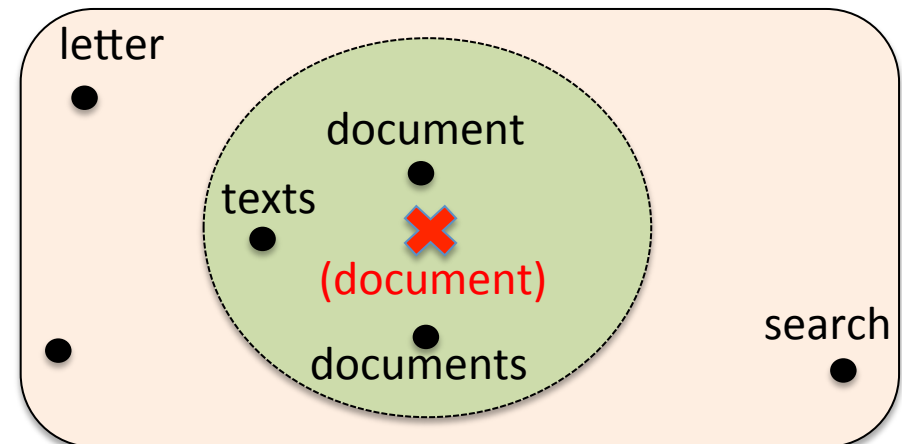
Observation

Synonyms are clustered in **word2vec** vector space.

Proposal

We introduce **latent synonym vectors** from which actual word vectors are generated with some noise.

In word vector space ...



✖ : synonym vector ● : word vector

光干渉断層計による網膜層厚からの緑内障患者の視野推定

○上坂俊允(東大), 森野佳生(東大), 村田博史(東大), 朝岡亮(東大), 山西健司(東大, CREST)

※ 本研究はJST-CRESTの一部として行われた

目標

網膜層厚からの
視野感度推定

背景

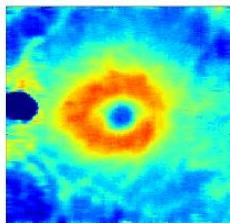
- ・ 視野感度の測定は臨床現場にとって負担
- ・ 近年, 光干渉断層計による網膜層厚計測が可能に

問題

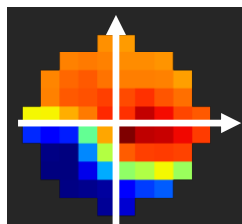
- ・ 同一患者眼の網膜層厚データと視野感度データの対は少ない

提案手法

網膜層厚データ



視野感度データ



1. 各ドメインで
特徴量を抽出

網膜層厚
特徴量

視野感度
特徴量

2. 同患者眼についてのデータ対から
潜在変数モデルを学習

マルチビュー学習

潜在
変数

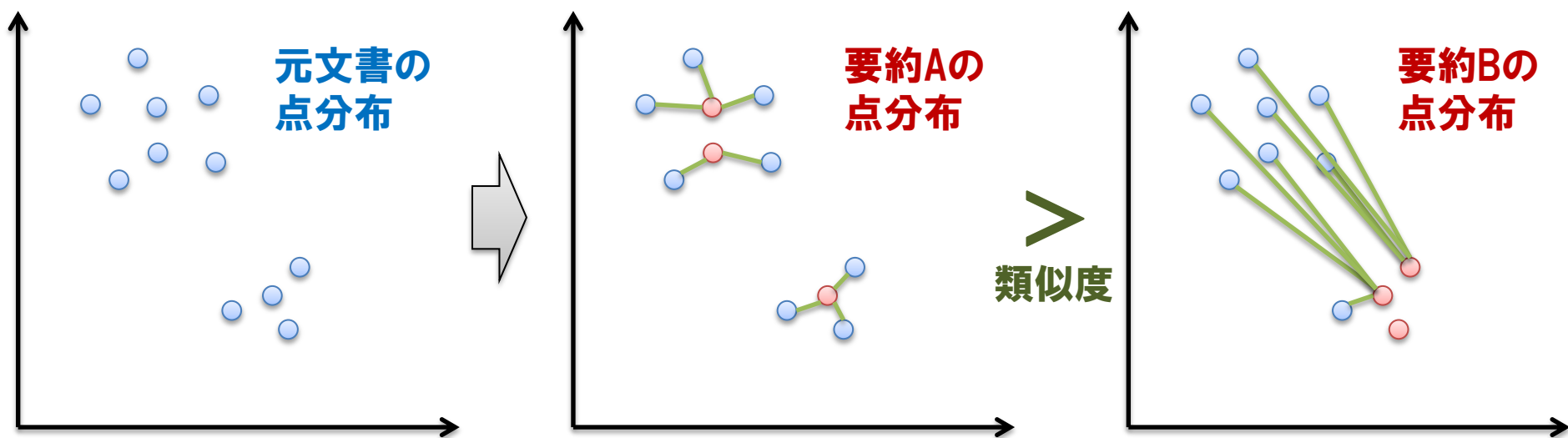
3. 学習したモデルにより
視野感度推定

分散表現に基づく

劣モジュラ最適化を用いた文書要約

小林隼人, 野口正樹, 谷塚太一 (ヤフー株式会社)

- 文書要約タスクを劣モジュラ最適化問題として定式化
- 分散表現に基づく新しい目的関数を提案
 - 文書を分散表現の点分布として表し, 元文書と要約の類似度を各点間の最近傍距離の和により算出
- 提案関数の単調劣モジュラ性を証明
 - KL情報量やEarth Mover's Distanceとも関係



再帰的紛失通信の提案と秘匿ゲノム検索への応用

清水 佳奈^{1,2}, 縫田 光司^{1,3}, Gunnar Rätsch²

1. 産業技術総合研究所, 2. Memorial Sloan Kettering Cancer Center, 3. JSTさきがけ

- ゲノム配列を秘匿したまま検索する手法を開発した。
- クエリの長さに対して線形の計算量と通信量で部分文字列/接頭辞の最長一致を検索することができる。
- 文字の置換や挿入削除を許す拡張などが可能。

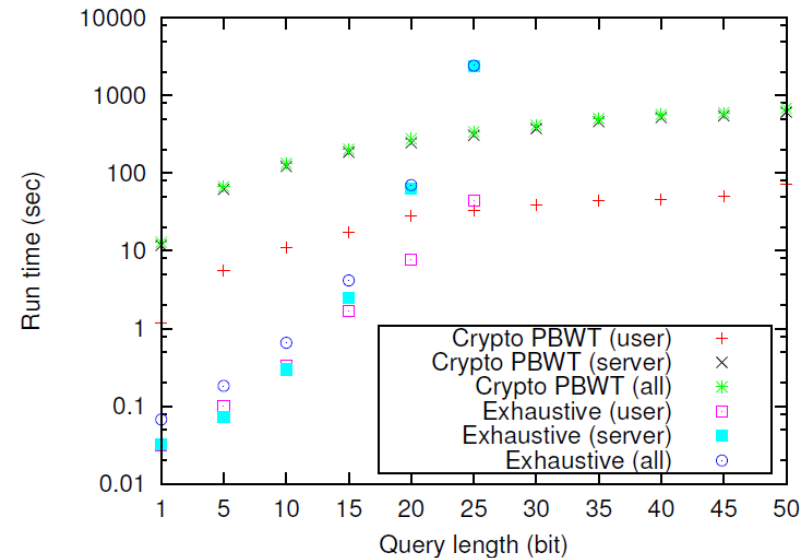
D-89

再帰的紛失通信

$f_k = v[f_{k-1}]$ とする。ユーザーが f_1 をサーバーに秘密にしたまま、サーバーのベクトル v から $f_k = v[\dots v[v[f_1]]\dots]$ のみを得ることができるアルゴリズムを考案した。

秘匿ゲノム検索

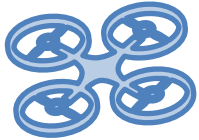
再帰的紛失通信を応用すると、ゲノム配列に効果的なデータ構造の(p)BWT上で、暗号化したまま文字列のマッチを計算することが可能。



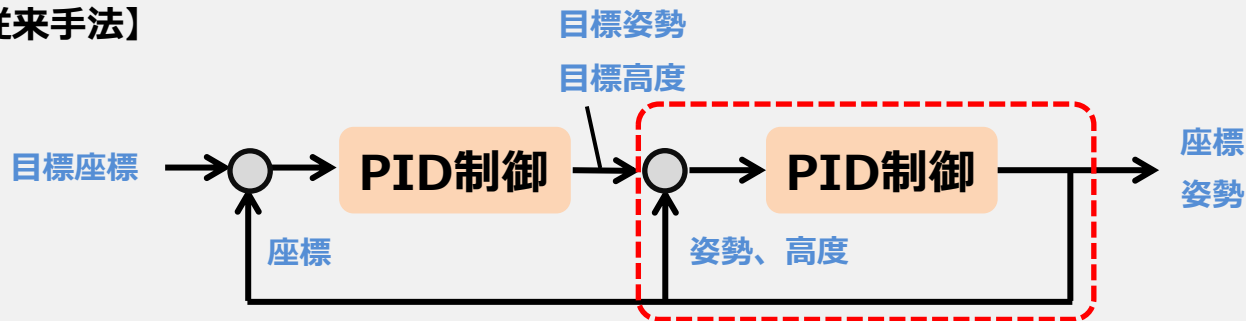
提案法 (crypto PBWT) と従来法 (Exhaustive) の比較. 1000ゲノムプロジェクトのデータを用いた。

江藤 力 (NEC)、鈴木 貴之 (NEC航空宇宙システム)、亀田 義男 (NEC)

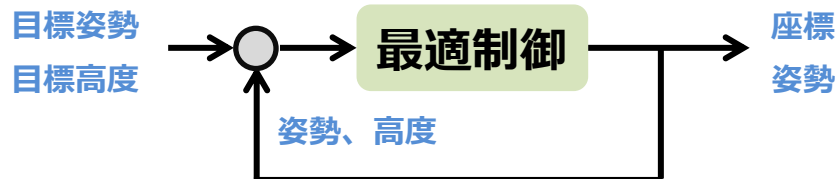
飛行制御 = (航路⇒姿勢) + (姿勢⇒ロータ制御)



【従来手法】



性能向上のため、
姿勢制御部分を
モデルベース制御へ！



【システム同定】

飛行データから
姿勢運動モデル
を学習

非線形ダイナミクスを
どう線形モデル化
するか？

D-93 動的シナプスを導入したリカレントニューラルネットワークを用いた時系列処理

「学生優秀プレゼンテーション賞対象」

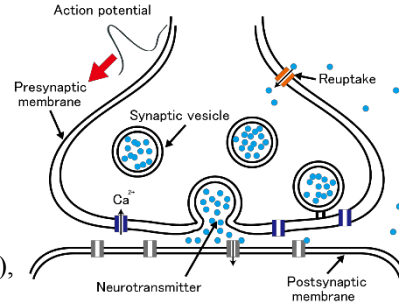
森 竜太 (東京大学情報理工学系研究科) ・ 香取 勇一 (公立ほこだて未来大学, 東京大学生産技術研究所)
 合原 一幸 (東京大学生産技術研究所)

動的シナプス

神経活動頻度に応じて神経伝達効率が一過的に変化するシナプス

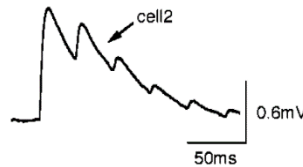
$$x_i(t+1) = x_i(t) + \frac{1-x_i(t)}{\tau_r} - z_i(t)x_i(t)u_i(t),$$

$$u_i(t+1) = u_i(t) + \frac{U_{se}-u_i(t)}{\tau_f} + U_{se}(1-u_i(t))s_i(t),$$

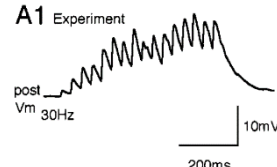


シナプス伝達効率: $x(t)u(t) / U_{se}$

Short-term depression (STD)



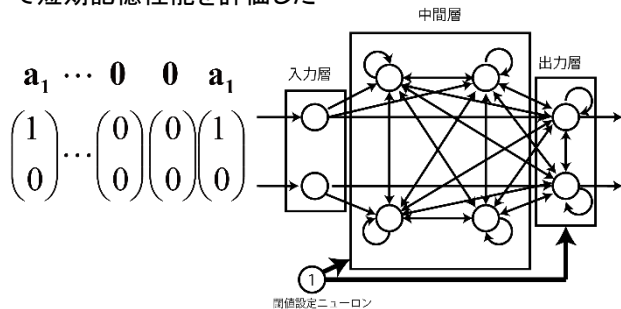
Short-term potentiation (STP)



(HENRY MARKRAM et al. 1998)

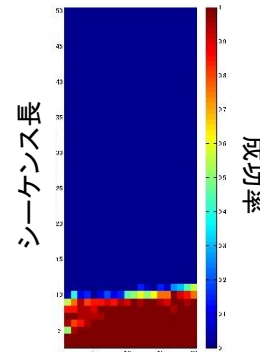
モデル

動的シナプスを導入した全結合型リカレントニューラルネットワークを用いて短期記憶性能を評価した



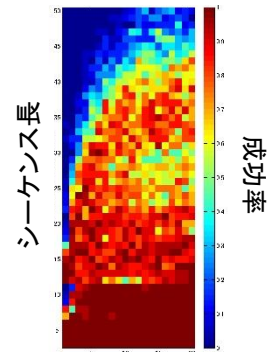
短期記憶課題

動的シナプスなし



中間層ニューロン数

動的シナプスあり

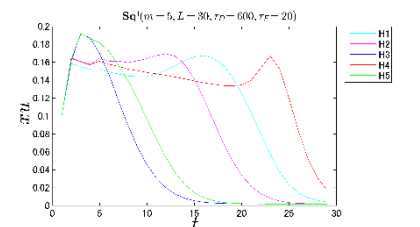
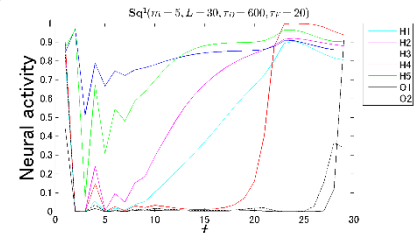
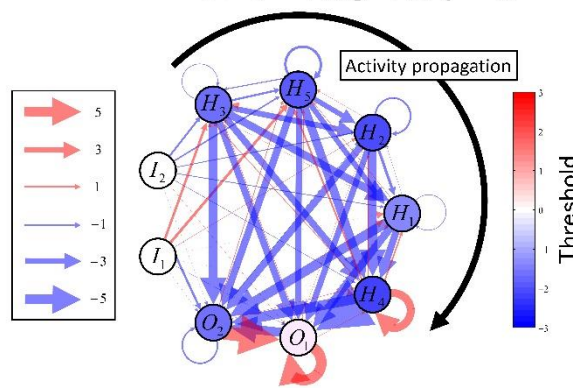


中間層ニューロン数

動的シナプスがリカレントニューラルネットワークの性能を大きく向上させることが分かった!

学習後のニューラルネットワーク構造

($m=5, L=30, \tau_D=600, \tau_F=20$)



目的

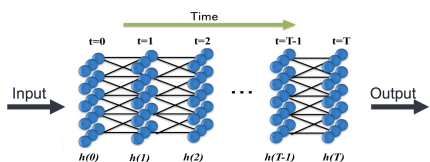
データの分布構造が、どのような幾何学的構造を持っているか知りたい。

仮説

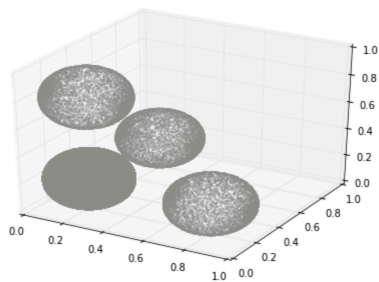
1. データは、それぞれの属するクラスに対応する低次元部分多様体に埋め込まれる。
 2. Deep Learningによって、その部分多様体を抽出できる。 ↑ [Refai, & Bengio et al., 2011]
- ⇒ これらが正しいければ、DNNの内部構造解析から、データの分布構造を知ることができる。

検証 Deep Learningの各種手法で、本当に多様体構造を学習できるか？

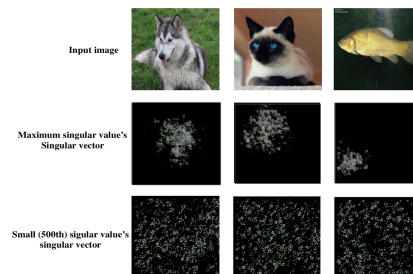
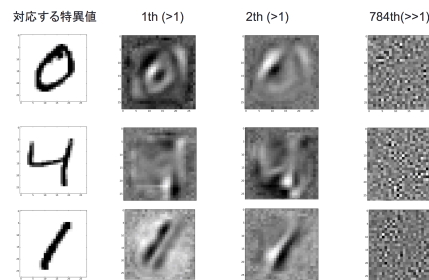
情報の圧縮・拡大を定量化



検証用データ



実データ



Jacobian matrix: representing the change of volume by variable transformation

$$J^{(t)} = \begin{pmatrix} \frac{\partial h_1^{(t)}}{\partial h_1^{(t-1)}} & \dots & \frac{\partial h_1^{(t)}}{\partial h_N^{(t-1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_N^{(t)}}{\partial h_1^{(t-1)}} & \dots & \frac{\partial h_N^{(t)}}{\partial h_N^{(t-1)}} \end{pmatrix}$$

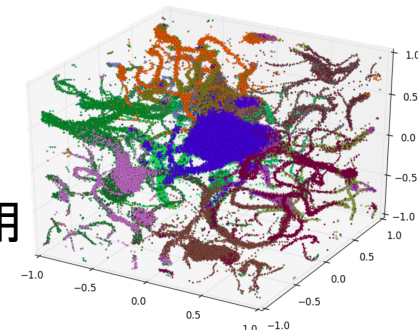
Whole jacobian $J = J^t \cdot J^{t-1} \cdot \dots \cdot J^0$

応用 > 大自由度力学系

Deep Learningで多様体構造が学習できるメリット

1. 大規模なデータセットに適用可能
2. DNNの内部解析で多様体の構造がわかる
3. 入力空間から多様体への変換関数を構成できる

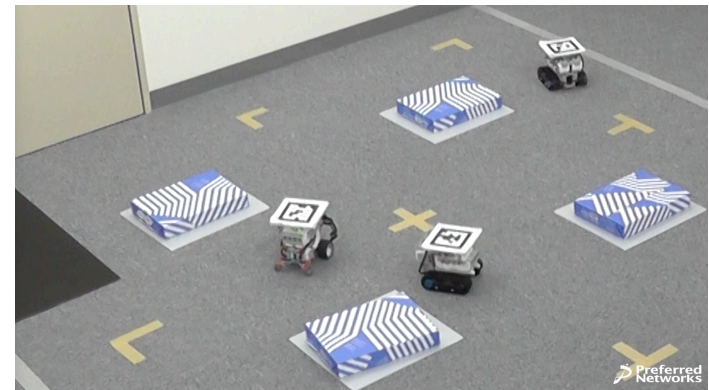
⇒ 力学系の縮約に応用



D-97: 深層強化学習によるロボット制御

松元 叡一(株式会社Preferred Networks)

- 深層学習と強化学習を用いてラジコンカーの運転操作を自動で学習する
 - 車の周囲の情報(障害物、目標コース)を273次元のベクトルで入力
 - 5通りの行動(前・後・左・右・ブレーキ)から各時刻で一つを選択
 - 行動価値関数をDeep-Q-Learningで学習
 - 複数台による分散学習とカリキュラム学習(衝突のペナルティを徐々に上昇させるなど)により2時間程度でスムーズな運動を獲得した



実機とシミュレーション双方で実験

