

第12回 情報論的学習理論ワークショップ ポスタープレビュー集

Collection of Preview Slides of
the 12th Workshop on Information-Based Induction
Sciences (IBIS 2009)



IBIS 2009

2009年10月19-21日
九州大学医学部 百年講堂

- 共催：
 - 九州大学グローバル COE プログラム「マス・フォア・インダストリ教育研究拠点」
 - 電子情報通信学会情報論的学習理論時限研究専門委員会
- 協賛：
 - 応用統計学会，システム制御情報学会，情報処理学会，情報理論とその応用学会（SITA），人工知能学会，IEICE 情報理論研究会，IEICE 人工知能と知識処理研究専門委員会，IEICE コンピューテーション研究会，IEICE ニューロコンピューティング研究専門委員会，IEICE パターン認識・メディア理解 (PRMU) 研究会，日本応用数理学会，日本行動計量学会，日本統計学会，日本ロボット学会，(社)計測自動制御学会，IEEE Committee Members of IT Society Japan Chapter，IEEE Computational Intelligence Society Japan Chapter
- 連続開催：データマイニングと統計数理研究会（DMSM 研究会）

巻頭言

情報論的学習理論ワークショップ（略称 IBIS; Information-Based Induction Sciences）は、情報理論、統計学、統計物理学、計算機科学など、広い意味で機械学習に関係する分野の学際理論フォーラムを作ることを目指し、1998年に産声をあげました。それ以来、参加者の方々の質の高い発表に支えられ、今年で12回目を迎えます。

第1回開催以来10年以上が経過しますが、その間社会の様相は大きく変わりました。インターネットを介した情報のやりとりが完全に社会基盤の一部となり、大量のデータから有用な情報を引き出す技術としてのデータマイニング、そしてその基礎を与える機械学習は、かつて半導体工学と量子力学がそうだったように、今や社会的に最も注目を浴びる学問分野になりつつあります。

第12回を迎える今年のIBISでは、これまでのIBISの伝統と、IBISをめぐる社会的・国際的な環境の変化に鑑み、次のようなフォーラムを提供することを目標にしました。

- ワールドクラスの質を備えた研究を生み出しうるという意味で国際的なフォーラムであること
- 同時に、研究者の出身分野や知識背景、所属組織にかかわらず、新しい萌芽的研究を迎え、生み出しうる開かれたフォーラムであること
- 理論的研究と社会的応用とをうまくバランスさせることで、機械学習に対する社会的期待に応えられるフォーラムであること

これらの目標に少しでも近づくため、IBIS 2009ではいくつかの新しい試みを行いました。第1に、「テクニカルトラック」と「ディスカッショントラック」という2トラック制で一般発表を募集したことです。第2に、4-8ページの投稿論文を集めた通常の予稿集に加えて、2ページのプレゼンテーション資料をまとめた概要集も公開することにしました。IBISが開かれたフォーラムとなるためには研究発表の内容を広く公開することが極めて重要ですが、これらはその用途に十分役に立つものと思います。

一般発表に加えて、IBIS 2009では、次の7つの企画セッションを開催します。

- 金融リスクと統計的学習
- 音声・音響処理と機械学習
- 化学構造とその数理
- 疎なネットワーク上のダイナミクス
- ランキング学習の最前線
- パターン認識の新潮流
- 広がる機械学習応用のフロンティア

金融、音声、化学、画像、制御、など、これらセッションに含まれる応用分野の広がり、先に述べたような機械学習の高い社会的需要を明瞭に表すものになっていると思います。また、コピュラ、極値統計、pre-image、経路積分、パタン符号化、など、これらのセッションに含まれる概念の多彩さは、機械学習研究の最近の急速な発展を象徴するものとなっていると思います。

機械学習という、おそらくすべての学問分野で最もホットな分野において、IBIS2009 という学際理論フォーラムが、研究上の新たな出会いと創造の一助になることを、実行委員一同期待しています。

2009年10月19日
IBIS 2009 実行委員会

第12回 情報論的学習理論ワークショップ 実行委員会

- 実行委員長：竹内純一（九州大学）
- プログラム委員長：井手剛（IBM）
- プログラム副委員長：中島伸一（ニコン）
- プログラム委員
 - － 猪口明博（大阪大学）
 - － 大羽成征（京都大学）
 - － 小山聡（北海道大学）
 - － 鹿島久嗣（東京大学）
 - － 加藤毅（お茶の水女子大学）
 - － 川喜田雅則（九州大学）
 - － 杉山将（東京工業大学）
 - － 瀬々潤（お茶の水女子大学）
 - － 中田貴之（NEC）
 - － 三村和史（広島市立大学）
 - － 持橋大地（NTT）
 - － 森本淳（ATR）
 - － 矢入健久（東京大学）
 - － 渡部晋治（NTT）

Preface

The IBIS Workshop is well established as a top-ranked research conference in machine learning in Japan. Since 1998 the IBIS Workshop has been a leading interdisciplinary forum where researchers and practitioners in various machine-learning-related disciplines can collaborate, including information theory, statistical science, statistical physics, computer science, data mining, and services sciences.

In the last decade of IBIS Workshops, we have seen drastic changes in many aspects of society: Massive information exchange through the Internet has become an essential part of the social infrastructure, and machine learning is receiving increasing attention from everyone as a basic and essential tool to extract useful knowledge from an enormous amount of data of the Internet age, just as quantum physics is the basis of semiconductor engineering.

The IBIS Workshop this year aims at being:

- International in promoting world-class research presentations from the Japanese community
- Open in encouraging new research attempts from new contributors and research domains
- Sound in supporting innovative research that matters in society while balancing theoretical and practical research work

Besides the technical program that covers all aspects of machine learning and related research, the IBIS Workshop this year features seven organized sessions, where top-notch researchers talk about the latest trends and activities in various research domains such as financial risk management, speech and audio processing, bioinformatics, network theories, machine learning of ranking, pattern recognition, and real-world applications in computer vision and plant monitoring.

To encourage discussions about ongoing studies also under submission to other conferences and journals, this year's IBIS Workshop accepted high-quality technical presentations on two tracks. The Technical Track accepted

4-8 page technical papers through a review process by the Program Committee, while submissions to the Discussion Track needed only an abstract in the form of presentation slides up to 2 pages. The accepted papers in the Technical Track are electronically published on the IBIS Web site as IBIS 2009 Technical Reports.

We hope the IBIS Workshop this year inspires your research by providing a high-quality technical program and well organized sessions with top-notch researchers, while also serving as an opportunity for networking between researchers from different areas.

Oct. 19, 2009
IBIS 2009 Organizing Committee

IBIS 2009 Committee

- Organizing Committee Chair: Jun'ichi Takeuchi (Kyushu University)
- Program Committee Chair: Tsuyoshi Idé (IBM)
- Program Committee Vice-Chair: Shinichi Nakajima (Nikon)
- Program Committee Members:
 - Akihiro Inokuchi (Osaka University)
 - Shigeyuki Oba (Kyoto University)
 - Satoshi Oyama (Hokkaido University)
 - Hisashi Kashima (the University of Tokyo)
 - Tsuyoshi Kato (Ochanomizu University)
 - Masanori Kawakita (Kyushu University)
 - Masashi Sugiyama (Tokyo Institute of Technology)
 - Jun Sese (Ochanomizu University)
 - Takayuki Nakata (NEC)
 - Kazushi Mimura (Hiroshima City University)
 - Daichi Mochihashi (NTT)
 - Jun Morimoto (ATR)
 - Takehisa Yairi (the University of Tokyo)
 - Shinji Watanabe (NTT)

目次

巻頭言	iii
Preface	vii
ポスターセッションA	1
テクニカルトラック	2
P007 代理ベイズ学習と隠れマルコフモデルへの応用 山崎 啓介(東京工業大学)	3
P008 カーネルマルコフ連鎖モンテカルロ法による測定誤差モデル推定 赤穂 昭太郎(産業技術総合研究所), 伊庭 幸人(統計数理研究所)	4
P009 一次元正規分布のなす空間への曲線あてはめ 藤木 淳(産業技術総合研究所), 赤穂 昭太郎(産業技術総合研究所)	5
P017 行と列の生成による線形計画ブースティング 畑埜 晃平(九州大学), 瀧本 英二(九州大学)	6
P025 Multiple Kernel Learning for Object Classification 中島 伸一((株)ニコン), Binder Alexander(Fraunhofer Institute FIRST), Müller Christina(Technische Universität Berlin), Wojcikiewicz Wojciech(Technische Universität Berlin), Marius Kloft(Technische Universität Berlin), Brefeld Ulf(Technische Universität Berlin), Müller Klaus-Robert(Technische Universität Berlin), Kawanabe Motoaki(Fraunhofer Institute FIRST)	7
P026 変分ベイズ法を用いた混合ベルヌーイ分布学習の相図について 梶 大介(東京工業大学), 渡辺 澄夫(東京工業大学)	8
P039 領域ベースの隠れ変数を用いた決定論的画像領域分割 三好 誠司(関西大学), 岡田 真人(東大, 理研)	9
P040 共起成分の含意関係に基づくデータマイニングの実験と考察 二木 克也(北海道大学), 湊 真一(北海道大学)	10
P044 High-Precision Speaker Verification by Adaptive Weighting of Local MFCC Features 坂井 俊亮(筑波大学), 亀山啓輔(筑波大学)	11
P048 オンライン学習可能な多重スケールでの時間発展を考慮したトピックモデル 岩田 具治(日本電信電話株式会社), 山田 武士(日本電信電話株式会社), 櫻井 保志(日本電信電話株式会社), 上田 修功(日本電信電話株式会社)	12
P051 Observational Reinforcement Learning Simm Jaak(東京工業大学), 杉山 将(東京工業大学), 八谷 大岳(東京工業大学)	13
P055 Matching between Piecewise Similar Curve Images 岩田 一貴(広島市立大学), 林 朗(広島市立大学)	14
P062 時系列パターンの多数決型識別器の設計 福富 正弘(九州大学), 小川原 光一(九州大学), 馮 堯楷(九州大学), 内田 誠一 (九州大学)	15
P066 変分ベイズ法における確定的アンニリングとハイパーパラメータの部分最適 化について 永田 賢二(東京大学), 片平 健太郎(科学技術振興機構), 岡ノ谷 一夫(科学技術 振興機構, 独立行政法人理化学研究所), 岡田 真人(東京大学大学院)	16
P071 Radon変換を介した医用画像再構成における画像修復 庄野 逸(電気通信大学), 岡田 真人(東京大学大学院)	17

P073	クラウドコンピューティングを用いた粒子フィルタのためのMapReduceアルゴリズム 石垣 司(産業技術総合研究所), 中村 和幸(明治大学), 本村 陽一(産業技術総合研究所)	18
P076	Virtual Concept Drift環境におけるRBFNNのモデル選択 山内 康一郎(中部大学)	19
P078	機械学習を用いたスプログ検出におけるHTML構造の類似性の利用 片山 太一(筑波大学), 芳中 隆幸(東京電機大学), 宇津呂 武仁(筑波大学), 河田 容英((株)ナビックス), 福原 知宏(東京大学)	20
P083	ベータ自由エネルギーとLoopy belief Propagation に現れるグラフのゼータ関数について 渡辺 有祐(総合研究大学院大学), 福水 健次(統計数理研究所, 総合研究大学院大学)	21
P089	劣モジュラカットとその応用 河原 吉伸(大阪大学), 永野 清仁(東京工業大学), 津田 宏治(産業技術総合研究所), Bilmes Jeff(Washington University)	22
P092	重み付きカーネルマシンの多次元パス追跡法に関する一考察 烏山 昌幸(名古屋工業大学大学院), 原田 尚幸(名古屋工業大学大学院), 竹内 一郎(名古屋工業大学大学院)	23
P099	ベイズ確率文脈自由文法のための高速構文木サンプリング法 武井俊祐(東京大学大学院), 牧野貴樹(東京大学), 高木利久(東京大学)	24
P104	VC Theory and a Concentration Inequality for Sums of Eigenvalues of Wishart Matrix 上野 康隆(東北大学大学院), 赤間 陽二(東北大学大学院)	25
ディスカッショントラック		26
P013	ベイズ推定と細胞の運命決定ダイナミクス 小林 徹也(東京大学)	27
P021	改良型双対尺度法を用いた医薬学データの数量化 山崎 広之(大阪大学), 岡本 晃典(大阪大学), 日高 伸之介(大阪大学大学院), 川下 理日人(大阪大学大学院), 高木 達也(大阪大学大学院)	28
P024	アイテム集合付きグラフからの協調関係抽出 福崎 睦美(お茶の水女子大学), 鹿島 久嗣(東京大学), 瀬々 潤(お茶の水女子大学)	29
P032	ベイズ推定を用いない曲指数型分布族の推定量の改善 川喜田 雅則(九州大学), 竹内 純一(九州大学)	30
P038	構造的欠損値を伴う多次元次系列データのための制約付き共埋め込み法 矢入 健久(東京大学)	31
P047	混合Bernoulli分布に基づく変分Bayes法による連想記憶モデルの解析 荒木 佑季(早稲田大学), 永田 賢二(東京大学), 岡田 真人(東京大学大学院), 井上 真郷(早稲田大学)	32
P061	定常Markovモデルの体積要素と拡大モデルについて 竹内 純一(九州大学)	33
P067	リスク考慮型強化学習に向けたリターン分布推定 森村 哲郎(日本IBM), 杉山 将(東京工業大学), 鹿島 久嗣(東京大学), 八谷大岳(東京工業大学), 田中 利幸(京都大学)	34
P084	SemiCCA: Efficient semi-supervised learning of canonical correlations 木村 昭悟(NTT), 亀岡 弘和(NTT), 杉山 将(東京工業大学), 前田 英作(NTT), 坂野 鋭(NTT), 石黒 勝彦(NTT)	35
P087	飽和性と極大性を用いた, 単一系列データにおける頻出系列の高速マイニング	36

	村田 拓也(山梨大学), 岩沼 宏治(山梨大学)	
P091	多次元尺度法を用いたアドレナリン受容体の機能分類 井上 泰仁(舞鶴工業高等専門学校)	37
P094	つながり方に着目した属性付きグラフの概要構造抽出 寺田 愛花(お茶の水女子大学), 瀬々 潤(お茶の水女子大学)	38
P097	教師なしアンサンブル適応法の提案と音響モデル適応への応用 篠崎 隆宏(東京工業大学), 久保田 雄(東京工業大学), 古井 貞熙(東京工業大学)	39
ポスターセッションB		40
テクニカルトラック		41
P004	1次元線形動的システムの特異性とベイズ汎化誤差への影響 内藤 卓人(東京工業大学), 山崎 啓介(東京工業大学)	42
P010	品質の異なる二つのデータ集合間の転移学習の解析 赤穂 昭太郎(産業技術総合研究所), 神嵐 敏弘(産業技術総合研究所)	43
P011	条件付きエントロピー最小化に基づく教師付き次元削減手法 日野 英逸(早稲田大学), 村田 昇(早稲田大学)	44
P033	ネットワーク科学の方法を用いたWebページネットワークの構造による分類 中川 帝人(名古屋大学), 鈴木 泰博(名古屋大学)	45
P037	化合物-タンパク質活性空間における特徴選択 新島 聡(京都大学大学院), 奥野 恭史(京都大学大学院)	46
P043	大幾何マージン最小分類誤り学習法 渡辺 秀行(独立行政法人 情報通信研究機構), 片桐 滋(同志社大学), 山田 幸太(同志社大学), マクダーモット エリック(日本電信電話株式会社), 中村 篤(日本電信電話株式会社), 渡部 晋治(NTT), 大崎 美穂(同志社大学)	47
P046	点過程を特徴付ける統計量の時間変化を推定する 下川 丈明(京都大学), 篠本 滋(京都大学)	48
P050	生物学情報への機械学習解析の応用(Toxicogenomicsへの展開) 武藤 裕紀(中外製薬株式会社), 松下 智哉(中外製薬株式会社), 芦原 基起(中外製薬株式会社)	49
P052	Ellipsoidal Support Vector Machines 門馬 道也(NEC)	50
P057	確率伝搬法による確率的画像処理における統計的性能評価 片岡 駿(東北大学), 安田 宗樹(東北大学), 田中 和之(東北大学)	51
P064	Chow-Liuアルゴリズムの一般化と、木の複雑さを考慮した修正版について 石田 悠(大阪大学大学院), 鈴木 譲(大阪大学大学院)	52
P070	Kullback-Leibler Importance Estimation Procedureを用いたRestricted Boltzmann Machineの学習アルゴリズム 桜井 哲治(東北大学), 安田 宗樹(東北大学), 田中 和之(東北大学)	53
P072	非定常2値時系列データから隠れた構造を読み取る - 神経科学データへの応用 - 瀧山 健(東京大学), 岡田 真人(東京大学大学院)	54
P075	部分空間SVMのための交差部分空間学習 井之上 直矢(東京工業大学)	55
P077	線形時間異種混合モデル選択のための期待情報量基準最小化法 藤巻 遼平(NEC), 森永 聡(NEC), 門馬 道也(NEC), 青木 健児(NEC), 中田 貴之(NEC)	56
P079	Extending the Use of Instrumental Variables for the Identification of Direct Causal Effects in SEMs	57

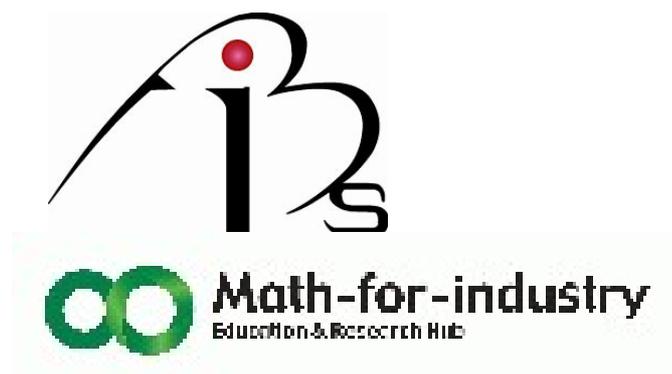
	Chan Hei(Center for Service Research, AIST), 黒木 学(大阪大学)	
P080	独立性の一般化に基づく統計モデルの拡張 藤本 悠(青山学院大学), 村田 昇(早稲田大学)	58
P082	近傍ハッシュを用いた高速なグラフカーネル 比戸 将平(IBM東京基礎研究所), 鹿島 久嗣(東京大学)	59
P086	Hannan-Quinn の命題は、線形回帰でも、ガウス型Bayesian ネットワークの構造推定でも正しい 鈴木 讓(大阪大学大学院)	60
P090	非線形プリコーディングの統計力学的解析 林 愛空(東京工業大学), 樺島 祥介(東京工業大学)	61
P095	Latent Dirichlet Allocationの量子アニーリング変分ベイズ学習 佐藤 一誠(東京大学), 栗原 賢一(Google東京), 田中 宗(東京大学), 宮下 精二(東京大学), 中川 裕志(東京大学)	62
P101	個性を考慮した周期的全身運動のオンライン予測 松原 崇充(奈良先端科学技術大学院大学), 玄 相昊(ATR), 森本 淳(ATR)	63
	ディスカッショントラック	64
P012	Sparse Exponential Family PCA with Heterogeneous Attributes 林 浩平(奈良先端科学技術大学院大学), 竹之内 高志(奈良先端科学技術大学院大学), 柴田 智広(奈良先端科学技術大学院大学), 池田 和司(奈良先端科学技術大学院大学)	65
P015	複数タスクに対するタスク選択実施能力を有するニューラルネットワーク 八尾 泰洋(東京大学大学院), 橋本 康弘(東京大学大学院), 陳 昱(東京大学大学院), 大橋 弘忠(東京大学大学院)	66
P023	How much is the sequence predictive for gene expression? 白石 友一(理化学研究所), 長嶋 剛史(理化学研究所), 岡田 眞里子(理化学研究所)	67
P027	各要素が混合ガウス分布に従う行列に対する行列因子化による欠損値予測 中村 政義(奈良先端科学技術大学院大学), 竹之内 高志(奈良先端科学技術大学院大学), 池田 和司(奈良先端科学技術大学院大学)	68
P034	階層状態無限隠れマルコフモデル 牧野 貴樹(東京大学)	69
P042	階層パターンを持つ自己相関型連想記憶モデルのPCAによる解析 塩塚 丁二郎(早稲田大学), 永田 賢二(東京大学), 岡田 真人(東京大学大学院), 井上 真郷(早稲田大学)	70
P049	ピアソングダイバージェンス最大化に基づく確率密度比推定のための次元削減手法 杉 山 将(東京工業大学), 原 聡(東京工業大学), Paul von Buenau(Technical University of Berlin), 鈴木 大慈(東京大学), 金森 敬文(名古屋大学), 川鍋 一晃(Fraunhofer FIRST.IDA)	71
P074	多様体正則化を用いたペアワイズ分類器の半教師付き学習 小山 聡(京都大学), 鹿島 久嗣(東京大学)	72
P081	候補系列抑制による高速系列データマイニング 水上 紘悠(山梨大学), 岩沼 宏治(山梨大学)	73
P085	A self-organizing state space approach to infer time-varying causalities between regulatory proteins 広瀬 修(東京大学), 清水 謙多郎(東京大学)	74
P088	多タスク学習のための二次錘計画法 加藤 毅(お茶の水女子大学), 鹿島 久嗣(東京大学), 杉山 将(東京工業大学), 浅井 潔(東京大学)	75
P093	Co-Clustering による協調フィルタリングと属性ベースフィルタリングの統合	76

- 麻生 英樹(産業技術総合研究所)
- P103 **スパース構造学習によるポットネット検知の検討** 77
- 村上 慎太郎(九州先端科学技術研究所, 九州大学), 濱崎 浩輝(九州先端科学技術研究所), 川喜田 雅則(九州大学), 竹内 純一(九州大学), 吉岡 克成(横浜国立大学), 井上 大介(情報通信研究機構), 衛藤 将史(情報通信研究機構), 中尾 康二(情報通信研究機構)
- P106 **隠れマルコフカーネルマシンを用いた系列データの識別とその音素認識タスクへの適用** 78
- 久保 陽太郎(早稲田大学), 渡部 晋治(NTT), マクダーモット エリック(日本電信電話株式会社), 中村 篤(日本電信電話株式会社), 小林 哲則(早稲田大学)

第12回情報論的学習理論ワークショップ (IBIS 2009)

2009年10月19日(月)-21日(水) 九州大学医学部 百年講堂

ポスターセッションA



第12回情報論的学習理論ワークショップ (IBIS 2009)

2009年10月19日(月)-21日(水) 九州大学医学部 百年講堂

ポスターセッションA

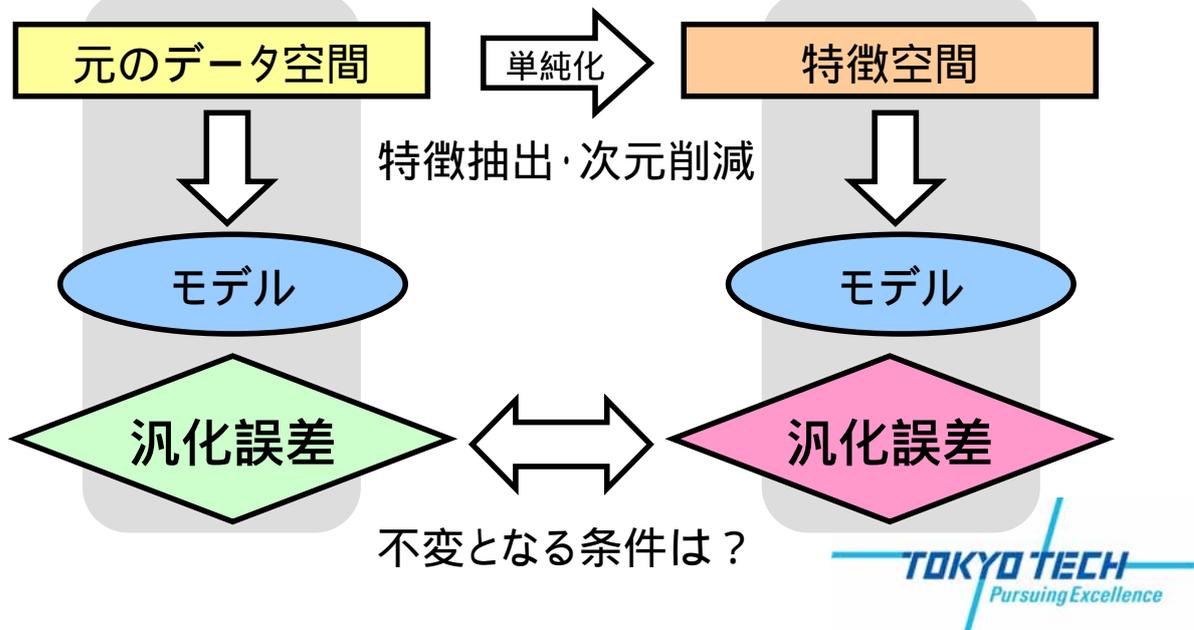
テクニカルトラック



P007

代理ベイズ学習と隠れマルコフモデルへの応用

山崎 啓介 (東工大) k-yam@pi.titech.ac.jp



P008

カーネルマルコフ連鎖モンテカルロ法 による測定誤差モデル推定

Measurement Error Model Estimation by
Kernel Markov Chain Monte Carlo Method

赤穂昭太郎

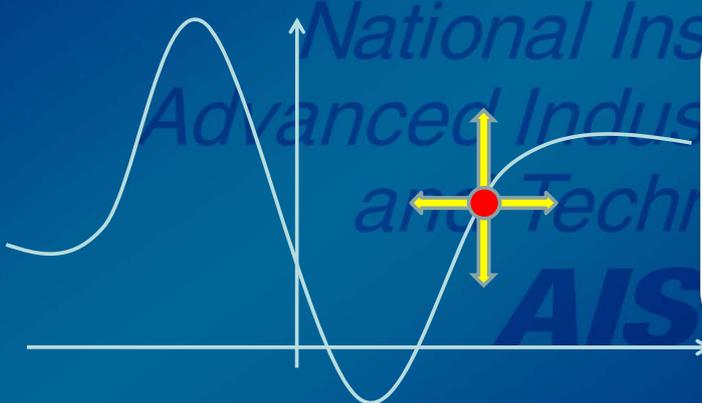
Shotaro Akaho

産総研 s.akaho@aist.go.jp

伊庭幸人

Yukito Iba

統数研 iba@ism.ac.jp



測定誤差モデル
 入力と出力の両方に
 ノイズがのる
 関数のモデルとして
カーネル(再生核)関数
 をとる

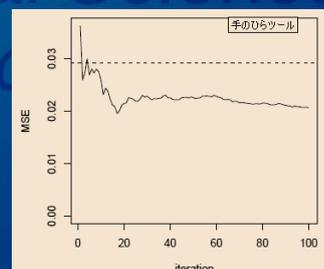
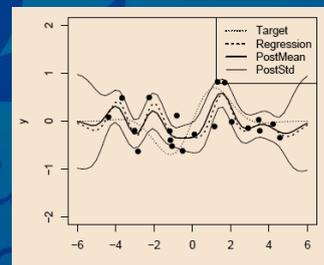
MCMCによるアプローチ

- 入りにノイズがあると潜在変数が無限の可能性
 →無限個の基底 (Representer 定理が使えない)
 →マルコフ連鎖モンテカルロ法 (MCMC) でシミュレーション

1. 関数のサンプリング
(無限次元)
2. 潜在変数をサンプリング

交換トリック

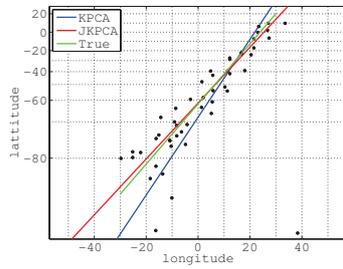
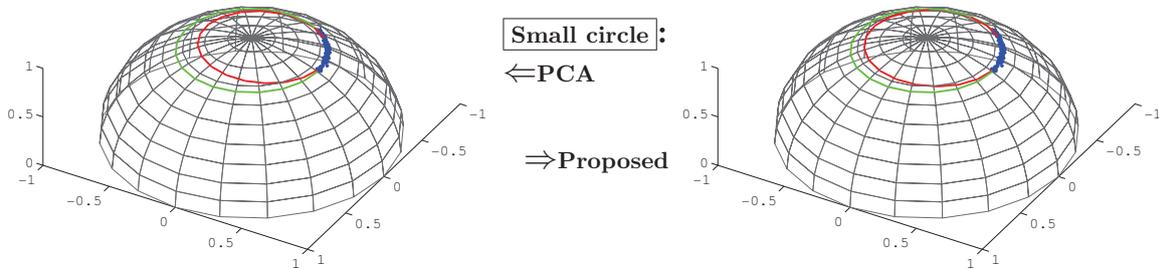
1. 潜在変数の候補を“先に”
サンプリング
2. 関数をサンプリング
(有限次元)
3. 潜在変数の採否を決める



P009:一次元正規分布のなす空間への曲線あてはめ

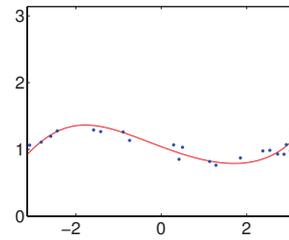
藤木淳・赤穂昭太郎（産総研）

動機：出来論（出来ちゃった結論）。意義決定，用途指導は範疇外。



← 航程線：メルカトル図法

ガウス核曲線：正距円筒図法 ⇒



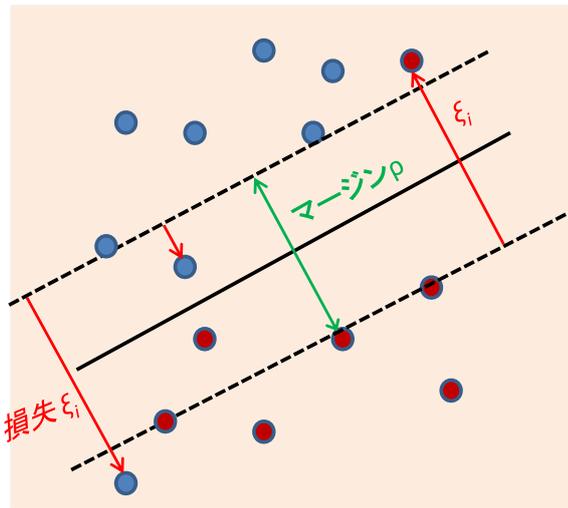
P17 行と列の生成による線形計画ブースティング

畑埜晃平・瀧本英二(九州大学)

1ノルムソフトマージン最適化

2ノルムソフトマージン最適化と並ぶ標準的な分類問題の定式化

「正例と負例をなるべく分離するマージンの大きな」超平面を求める



$$\max \rho - \nu \sum_{i=1}^m \xi_i \quad (\nu \text{ は定数}, \nu \geq 1)$$

sub.to.

$$y_i \left(\sum_{j=1}^n \alpha_j h_j(\mathbf{x}_i) \right) \geq \rho - \xi_i, \quad (i = 1, \dots, m, y_i = 1, -1)$$

$$\|\alpha\|_1 = \sum_{i=1}^n |\alpha_i| = 1, \alpha_i \geq 0 \quad \text{1ノルムで正規化}$$

特徴

- 線形計画問題
- 最適解が疎になりやすい(特徴選択に有効)

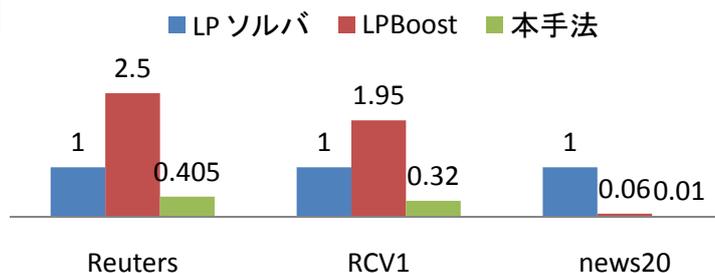
提案手法: Sparse LPBoost

- 与えられた $\epsilon > 0$ に対して, 1ノルムソフトマージン最適化問題の ϵ 近似解を出力
- 元の問題を複数の(よりサイズの小さい)線形計画問題に分解して繰り返し解く
- 最適解の疎性を利用し高速化

実験結果

データセット	例の個数	仮説の個数
Reuters-21578	10,170	30,389
RCV1	20,242	47,237
news20	19,996	1,355,193

実行時間比



ポスター番号：25

Multiple Kernel Learning for Object Classification

Shinichi Nakajima, Alexander Binder, Christina Müller, Wojciech Wojcikiewicz,
Marius Kloft, Ulf Brefeld, Klaus-Robert Müller, Motoaki Kawanabe

Nikon, Fraunhofer Institute First, Technische Universität Berlin

画像識別



→ 椅子



→ 人、羊

$$\{k_l(x, x')\}_{l=1}^L$$

SIFT, phog, ...
grey, RGB, opponent color, ...
pyramid structure

複数の特徴量（カーネル）を組み合わせるが、
従来法（sparseMKL）では精度が上がらない場合が多い。

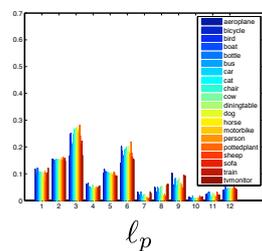
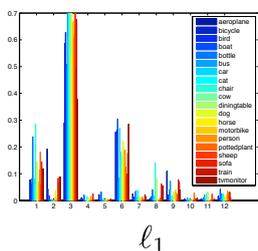
⇒ non-sparse MKL

$$K(x, x') = \sum_l \beta_l k_l(x, x') \quad \{\beta_l \geq 0\}$$

~~$$\sum_l \beta_l = 1$$~~

non-sparse
constraint

$$\sum_l \beta_l^p = 1 \quad (p \geq 1)$$



	uniform	ℓ^1	ℓ^p -joint	ℓ^p -single
mean AP	40.8±1.0	40.8±0.9	42.6±0.7	42.3±0.9

P26 変分ベイズ法を用いた混合ベルヌーイ分布学習の相図について

梶 大介 渡辺 澄夫 [東京工業大学]

【内容】変分ベイズ学習におけるハイパーパラメータの影響

混合ベルヌーイ分布を用いて、混合比/混合分布側両方のハイパーパラメータが、相転移点に与える影響について実験的に調査した。



【方法】

ハイパーパラメータを変化させた場合の予測分布の混合比を調べ、相図を作成。

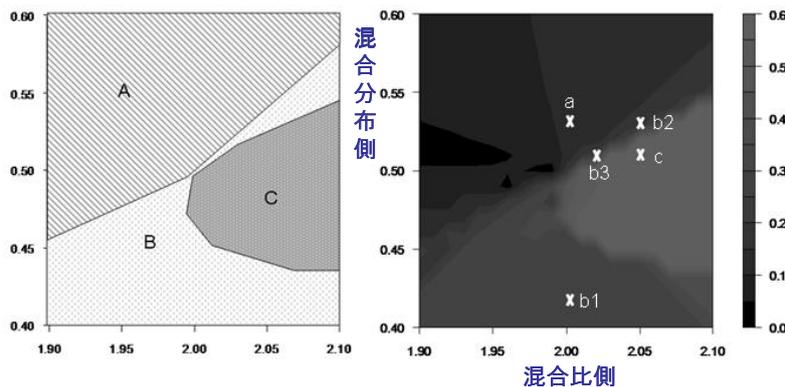


【分かったこと】

相転移点は混合比/混合分布側両方のハイパーパラメータに依存して変化し、予測分布の挙動は、大きく3つのパターンに分類できる。

結果

【混合分布側の次元が3の場合の相図】

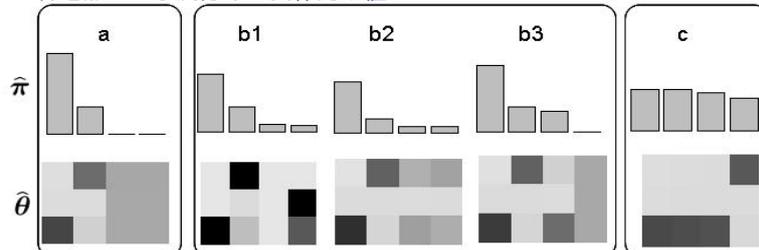


予測分布の挙動

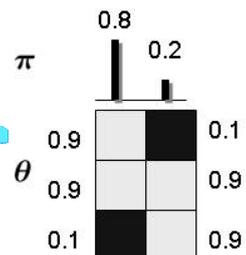


混合比側および混合分布側双方のハイパーパラメータに依存

各地点での予測分布の具体的な値



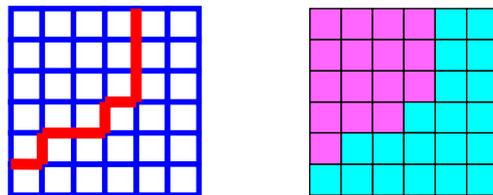
真の分布



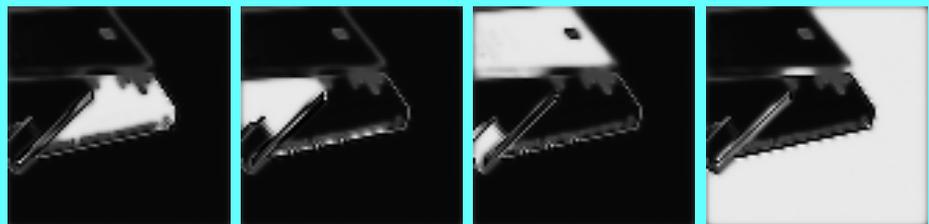
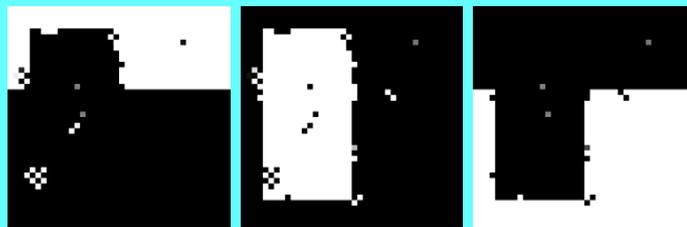
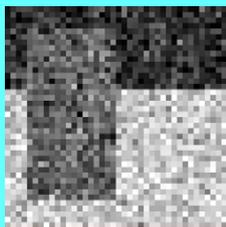
39 領域ベースの隠れ変数を用いた 決定論的画像領域分割

三好 誠司(関西大) 岡田 真人(東大, 理研)

- MRFとベイズ推定に基づく画像領域分割
- ホッツスピンの型・領域ベースの隠れ変数
- 隣接画素値情報の有効活用
- 変分推論による決定論的アルゴリズム



39 領域ベースの隠れ変数を用いた 決定論的画像領域分割



P040 共起成分の含意関係に基づく データマイニングの実験と考察

二木 克也, 湊 真一

目標 アイテムの組み合わせの集合であるデータベースより、従来手法では、発見できなかった意味のある関係を抽出する。

共起成分の含意関係 → 直接的共起関係に基づく手法では、発見できない関係が抽出できる

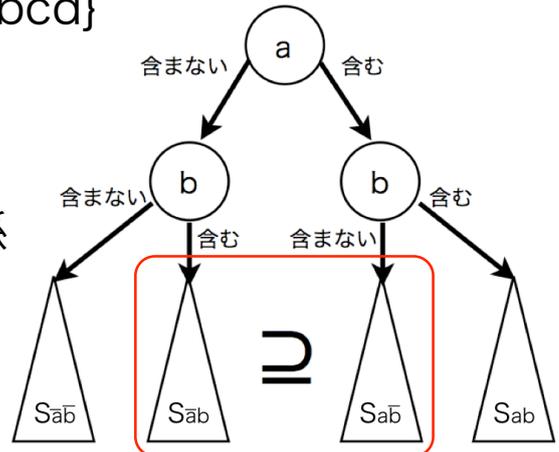
例: $S = \{bc, acd, ce, ac, abe, bd, bcd\}$

$$\left. \begin{matrix} S_{\bar{a}b} = \{c, d, cd\} \\ S_{ab} = \{cd, c\} \end{matrix} \right\} S_{\bar{a}b} \supseteq S_{ab}$$

aはbに対して共起成分の含意関係

全てのアイテムの組について実行

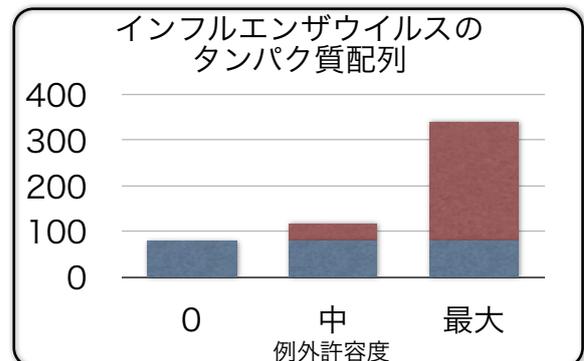
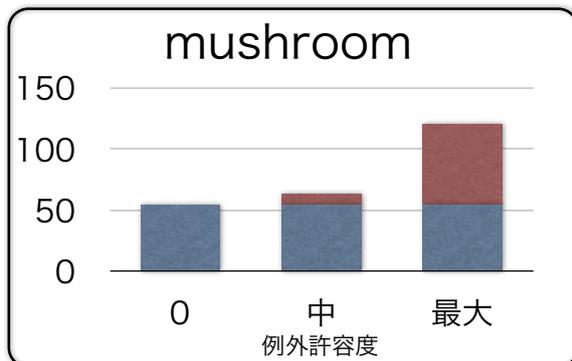
今まで知られていない 有意な関係を抽出 **アイテム数の二乗回 → ZDDで高速化**



ところで、従来手法では、共通成分がどれほど多くても、例外が1件でもあれば、抽出できなかった

例外を許容してみる

実験結果



例外を一定の水準まで許容した場合
意味のある新たな関係が抽出された

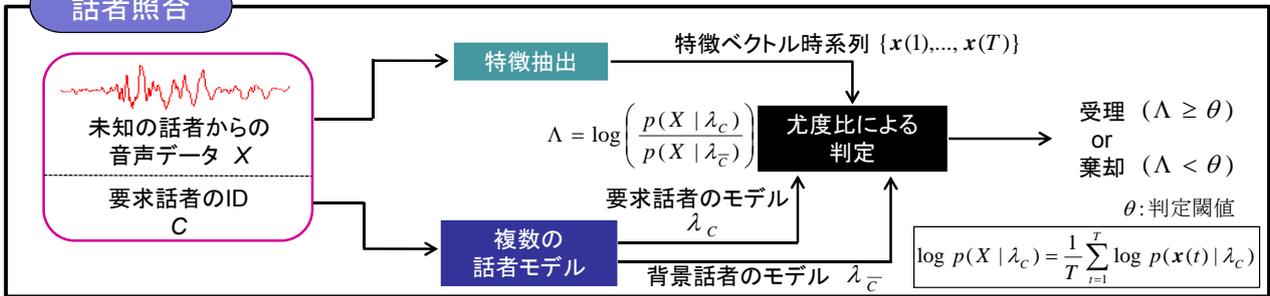
例外を最大限許容した場合
n者択一型のデータベースに関しては
設問ごとのアイテムの集合が抽出された

■ 例外を許した場合
■ 例外を認めない場合

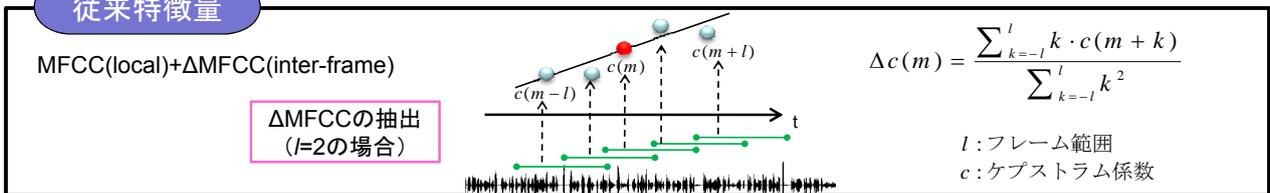
44 局所MFCC特徴量の最適線形結合による テキスト独立型話者照合

坂井俊亮 亀山啓輔
筑波大学 システム情報工学研究科

話者照合



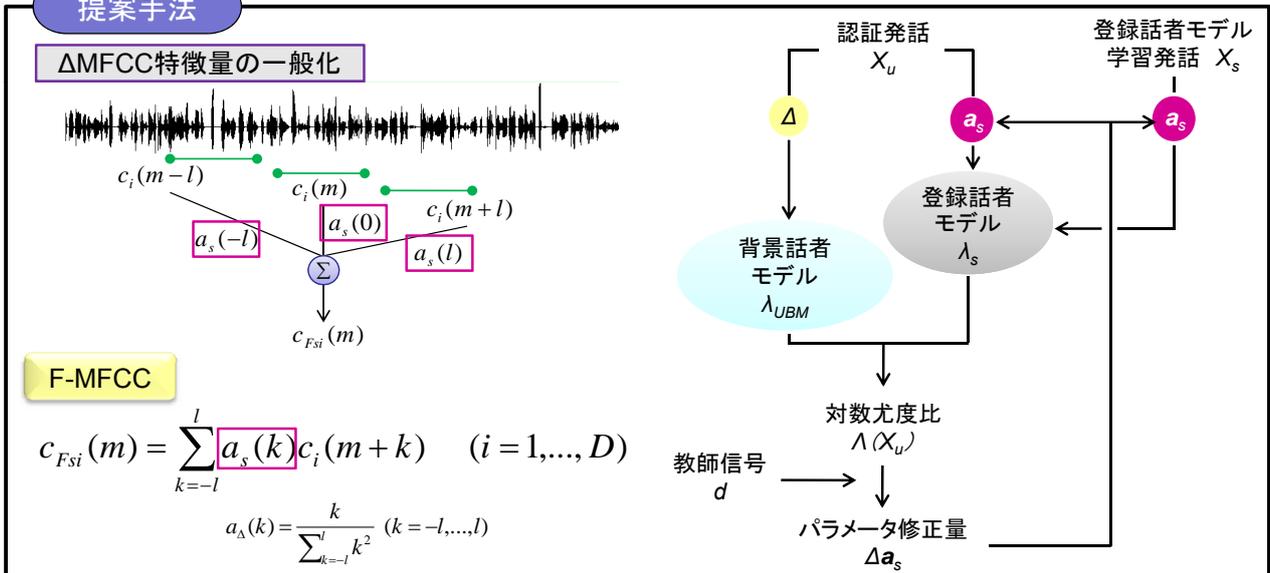
従来特徴量



問題点

- 単一フレーム特徴量のみでは十分な精度が得られない
- ΔMFCC特徴量におけるフレーム間の係数パラメータ $a(k)$ が固定

提案手法



実験結果

EER (%)

	∞dB	30dB	20dB
MFCC	2.80	2.60	2.20
MFCC+ΔMFCC	2.69	1.94	2.05
MFCC+F-MFCC	2.20	1.89	1.34

オンライン学習可能な多重スケールでの 時間発展を考慮したトピックモデル

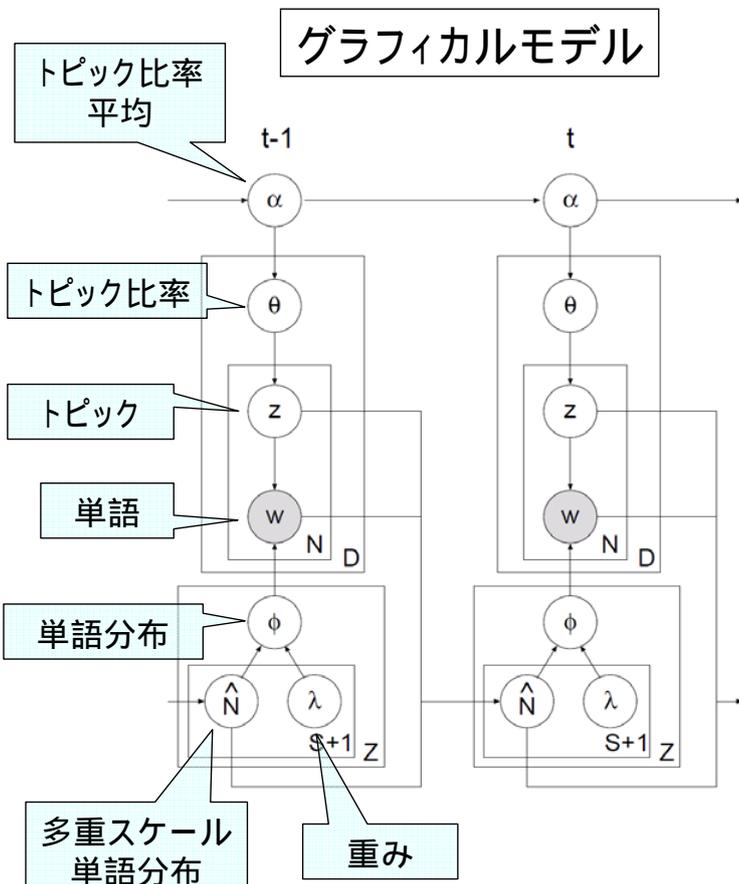
岩田具治, 山田武士, 櫻井保志, 上田修功
NTTコミュニケーション科学基礎研究所

目的

– 日々蓄積される文書データを逐次的に解析

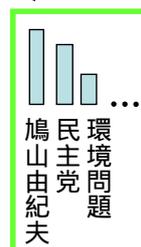
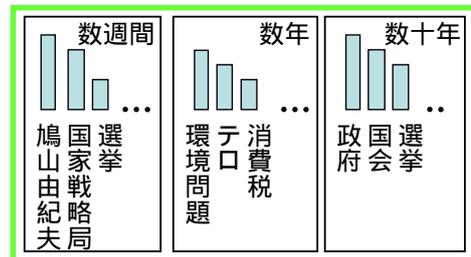
特徴

- 複数の時間スケールの時間発展を考慮することにより, **頑健なモデル推定**が可能
- オンライン学習により大規模ストリームデータでも**高速な学習**が可能



事前単語確率は
一時刻前の多重スケール
単語確率の重み付き和を
平均とするディリクレ分布

時刻 $t-1$ の複数の時間スケール
での単語確率



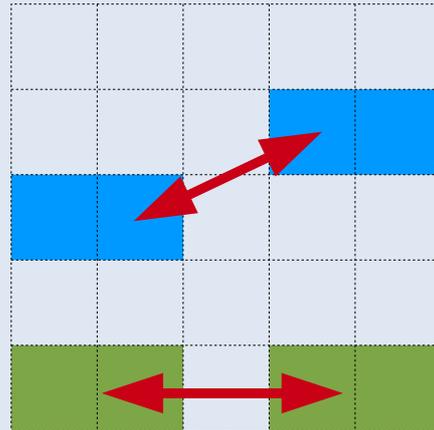
例: 政治トピック

時刻 t の単語確率

Observational Reinforcement Learning

Simm, Sugiyama, Hachiya Tokyo Inst. of Technology

- **Goal of RL:** finding the policy maximizing the rewards in **unknown** environments.
- **Current shortcoming:** **too many** samples are needed to find good policies.
- **Proposed idea:** share data between **similar** regions based on additional **observations**.



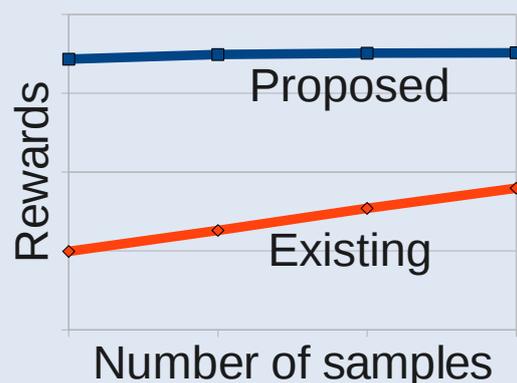
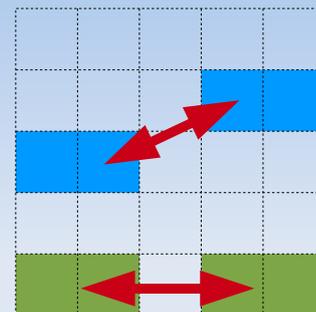
Observational Reinforcement Learning

Simm, Sugiyama, Hachiya Tokyo Inst. of Technology

- **Solution:** apply **multitask learning** to similar regions.
- Application of ORL to object lifting task



- **Come to my poster!**



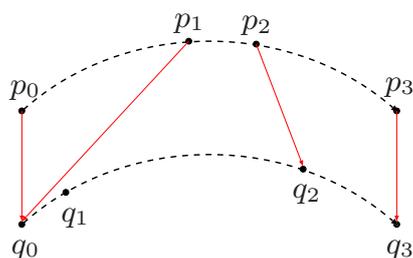
55: Matching between Piecewise Similar Curve Images

岩田 一貴 and 林 朗 (広島市立大学)

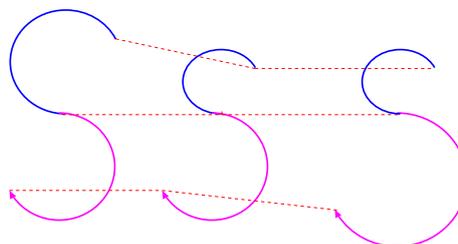
□ 曲線整合: 2つの曲線像上の標本点の間の対応を取ることで、一般に形状解析は曲線整合を基にして行われる。

□ 区分的な相似 (Piecewise Similarity): 人間が形を認識する際には区分的に相似な形は同じ形に見える。区分的な相似は同値関係。

👉 ポスター発表では、区分的に相似な曲線像のための曲線整合アルゴリズムを提案

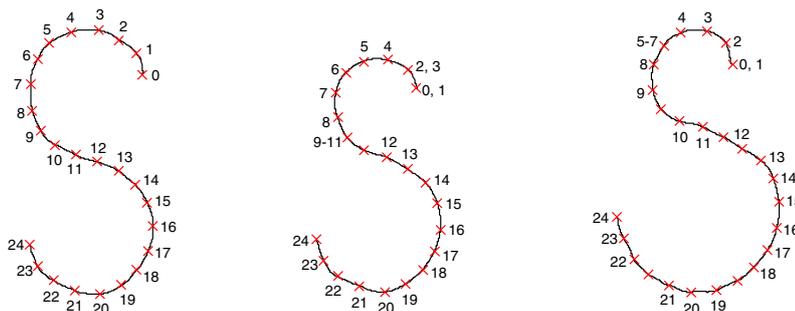


(a) 曲線整合



(b) 区分的に相似な曲線像

□ 実験結果:



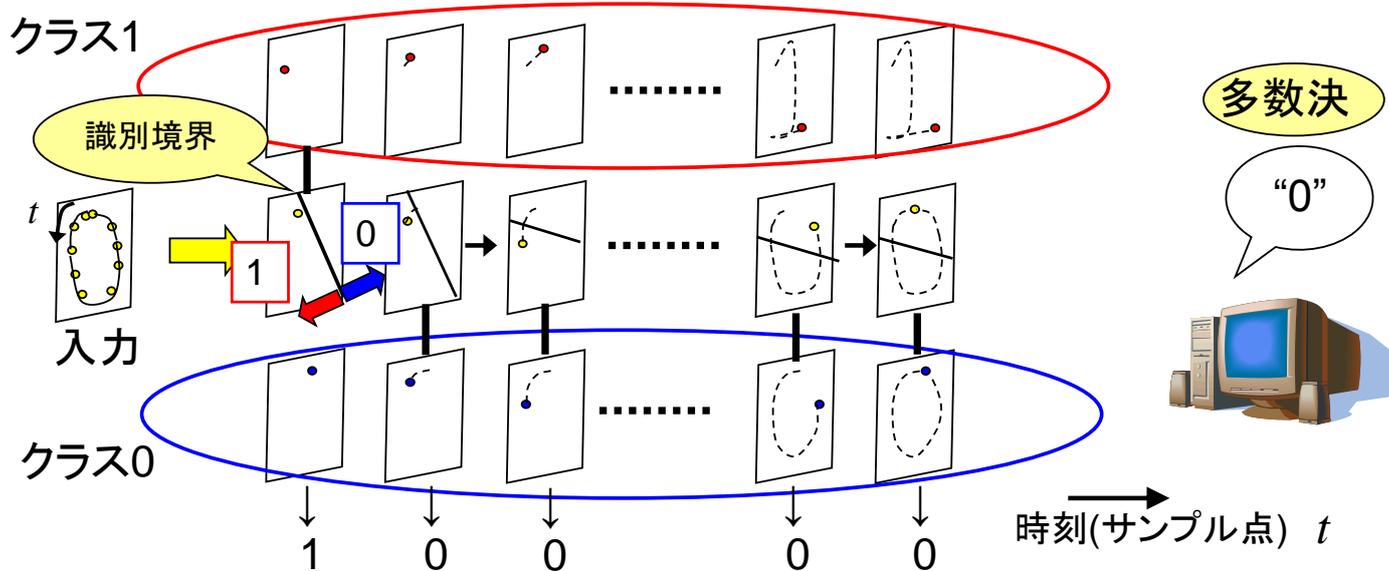
□ アルゴリズムの特徴:

- 曲線整合と曲線のサンプリングを同時に考える
- 単純で実装が簡単
- 得られる整合に関してある理論的な保証がある
- 相似な関係にある部分曲線の両端の点とそのスケールが事前にわからなくても良い

62 時系列パターンの多数決型識別器の設計

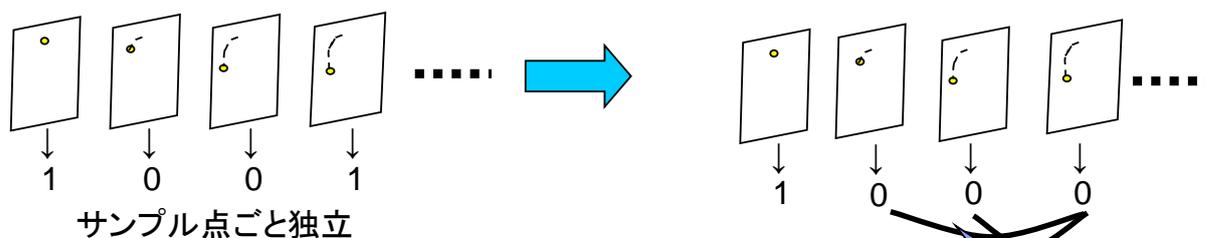
福富正弘, 小川原光一, 馮堯鏜, 内田誠一 (九州大学)

■ 各時刻での識別結果の多数決によりクラスを決定



Key points

- サンプル点对のクラスラベルが異なる場合にペナルティを導入
 - サンプル点对のクラスラベルの変化を小さく
 - サンプル点对に相互制約



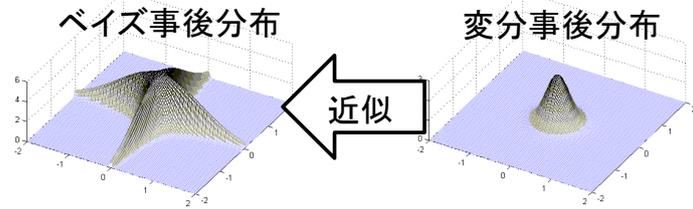
- 学習により適切なサンプル点对に導入
 - 誤認識の減少
- グラフカットによりラベルを割り当て

P066 変分ベイズ法における確定的アニーリングとハイパーパラメータの部分最適化について

永田賢二¹, 片平健太郎^{1,2}, 岡ノ谷一夫^{2,3}, 岡田真人^{1,3}
 1,東京大学 2,JST ERATO 3,理研 脳科学

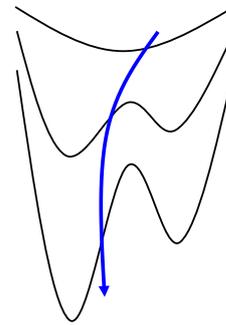
変分ベイズ法(Attias, 1999)

- ・ベイズ学習の近似手法の一つ
- ・局所最適解の問題を有する。



変分ベイズにおけるアニーリング(Katahira et. al., 2007)

目的関数に逆温度パラメータを導入することで局所解を避ける。



本研究では、

変分ベイズ法における新たなアニーリング法の提案
 (確定的アニーリング) + (ハイパーパラメータの部分最適化)

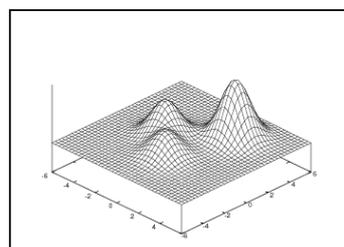
$$\text{ベイズ事後分布: } p(Y, \theta | X) \propto p(X, Y | \theta) p(\theta)$$

(尤度関数) × (事前分布) により構成

$$\text{従来のアニーリング: } p(Y, \theta | X, \beta) \propto \{p(X, Y | \theta) p(\theta)\}^\beta$$

$$\text{提案手法: } p(Y, \theta | X, \beta_1, \beta_2) \propto p(X, Y | \theta)^{\beta_1} p(\theta)^{\beta_2}$$

- ・混合正規分布モデルにおいて、有効性を実験的に示す。

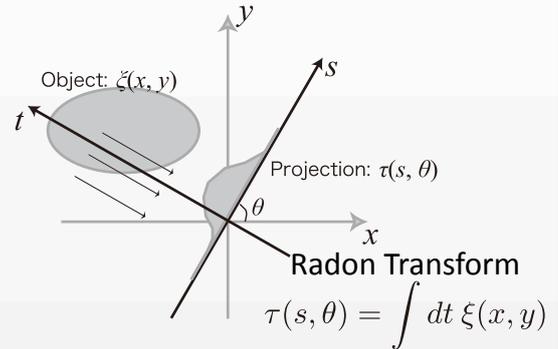


P071 Radon変換を介した医用画像再構成における画像修復

庄野 逸(電気通信大学 情報通信工学科), 岡田真人(東京大学 大学院新領域創成科学研究科)

shouno@ice.uec.ac.jp

- PET/CT 画像における画像修復を取り扱う。撮像過程は Radon 変換と呼ばれる積分変換で定式化し、これにノイズが乗ることを仮定して画像修復を行う。



- 画像修復は Bayes 推定に基づいて行う。

- 観測過程: Radon変換 + ガウスノイズ

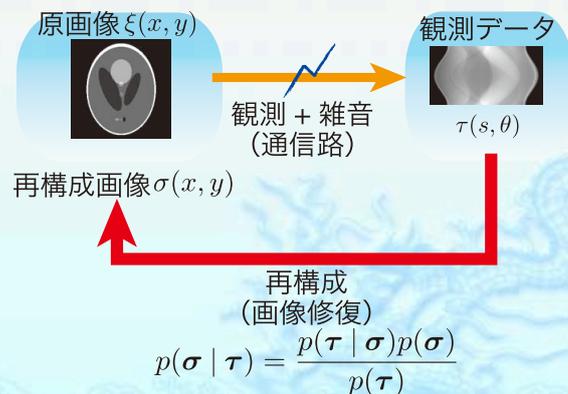
$$H_n(\tau | \sigma) = 4\pi^2 \gamma \int ds \int d\theta \left(\tau(s, \theta) - \int dt \sigma(x, y) \right)^2$$

$$p(\tau | \sigma) \propto \exp(-H_n(\tau | \sigma))$$

- 事前分布: Markov Random Field 状の分布

$$H_{\text{pri}}(\sigma) = \beta \iint dx dy \|\nabla \sigma\|^2 + 4\pi^2 h \iint dx dy \|\sigma\|^2$$

$$p(\sigma) \propto \exp(-H_{\text{pri}}(\sigma))$$



IBIS2009 @ Kyushu Univ. 2009/10/19~21

ハイパーパラメータ推定を用いた画像修復結果

- 事後分布による画像修復

- Fourier 領域では可解モデルとして取り扱うことが可能。

$$p(\sigma | \tau) \propto \exp\left(-4\pi^2 \int d\theta \int d\tilde{s} \tilde{F} \|\tilde{\sigma} - \frac{\gamma}{\tilde{F}} \tilde{\tau}\|^2\right)$$

$$\times \exp\left(-4\pi^2 \int d\theta \int d\tilde{s} \gamma \left(1 - \frac{\gamma}{\tilde{F}}\right) \|\tilde{\tau}\|^2\right)$$

where $\tilde{F} = (\beta \tilde{s}^2 + h) |\tilde{s}| + \gamma$

Reconstructed Image (in Fourier domain) $\langle \tilde{\sigma}_{\tilde{s}, \theta} \rangle = \frac{\gamma}{\tilde{F}} \tilde{\tau}_{\tilde{s}, \theta}$

- ハイパーパラメータ β, h, γ の推定

- 対数周辺事後分布 $\ln p(\tau; \beta, h, \gamma)$ の最大化により求める。

$$\ln p(\tau; \beta, h, \gamma) = \ln Z_{\text{post}}(\beta, h, \gamma) - \ln Z_{\text{pri}}(\beta, h) - \ln Z_n(\gamma)$$

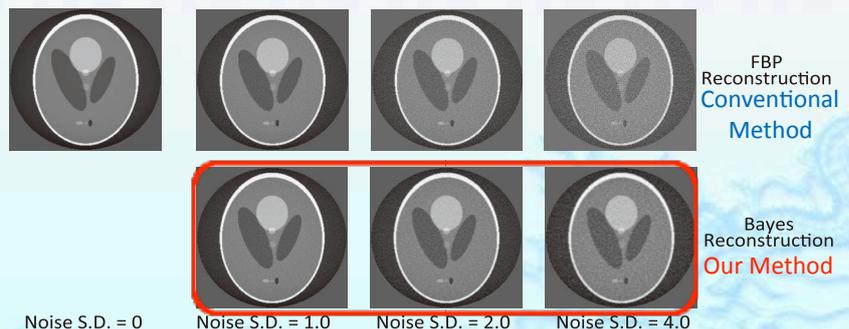
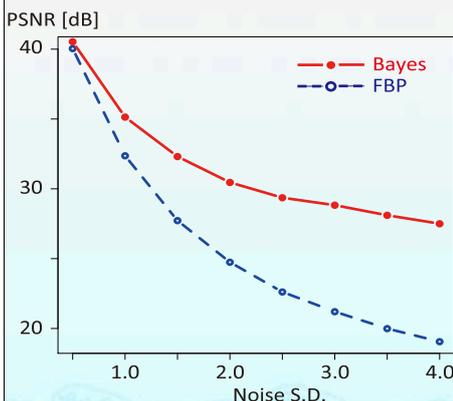
where

$$Z_{\text{pri}}(\beta, h) = \sum_{\sigma} e^{-H_{\text{pri}}(\sigma; \beta, h)}$$

$$Z_n(\gamma) = \sum_{\tau} e^{-H_n(\tau | \sigma; \gamma)}$$

$$Z_{\text{post}}(\beta, h, \gamma) = \sum_{\sigma} e^{-H_{\text{pri}}(\sigma; \beta, h) - H_n(\tau | \sigma; \gamma)}$$

- 目視と PSNR による画質評価を行い、従来手法からの画質向上を確認した。



IBIS2009 @ Kyushu Univ. 2009/10/19~21

P073 クラウドコンピューティングを用いた 粒子フィルタのためのMapReduceアルゴリズム

石垣 司¹、 中村和幸²、 本村陽一¹

所属1：産総研サービス工学研究センター 所属2：明治大学研究・知財戦略機構

研究対象

- ・大規模データ処理や大規模モデル計算の重要性は高まっており、それに応じて並列分散処理の必要性も高まっている。しかし、並列分散処理に関するリテラシーは情報科学・計算科学以外の分野では高くない。
- ・MapReduceと呼ばれる並列分散処理を簡易に行うプログラミングモデルが提案され、クラウドコンピューティング上での使用が可能となっている。
- ・粒子フィルタのMapReduceフレームワークによる実現と並列化計算の実験

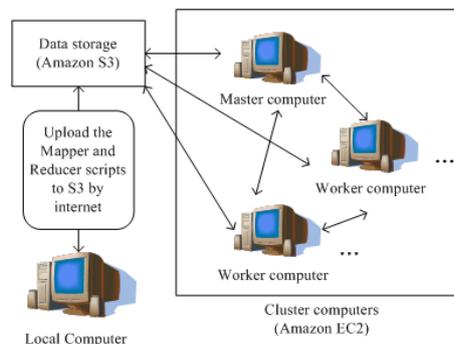


実験

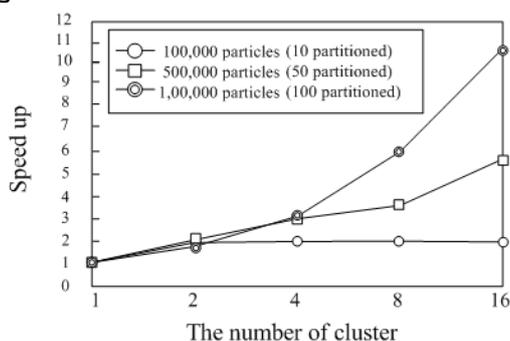
- ・10万粒子、50万粒子、100万粒子を使用
- ・クラスタ数は1, 2, 4, 8, 16
- ・各粒子は1セット1万粒子に分割
- ・Amazon Elastic MapReduceを利用

実験結果

- ・100万粒子使用時に並列化の効果を確認
- ・16クラスタ使用時には計算時間が5.5時間から31分に短縮
- ・並列分散処理の実行環境が数十分で実現必要な作業（初回のみ）
 - －Amazonへの登録（約15分）
 - －クラウド上の環境整備（約5分）
 - －ローカルPCの環境整備（約15分）



Amazon Elastic MapReduce サービス概要



実験結果（計算速度の倍率）



Virtual Concept Drift環境におけるRBFNNのモデル選択 (76)

中部大学工学部 情報工学科 山内康一郎

Virtual Concept Drift

学習サンプルの分布のpriorが変動

$$p(\mathbf{x}, y) = p(y | \mathbf{x}) p(\mathbf{x})$$

変動

追記学習環境と仮定する

追記学習アルゴリズム精緻化

共変量シフトの一種とみなしてモデル化

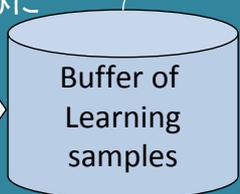


この環境に適したRBFNNの追記学習アルゴリズムの導出、およびモデル選択規準を導く。

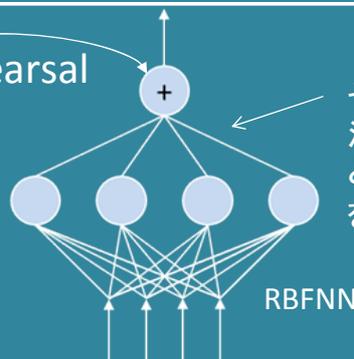
前提

- 逐次的にサンプルが来る
- 新しいサンプルと出会うたびにバッファの中身を全て学習

(\mathbf{x}_p, y_p)



rehearsal



その時点で汎化誤差最小となるRBFNNを構築したい。

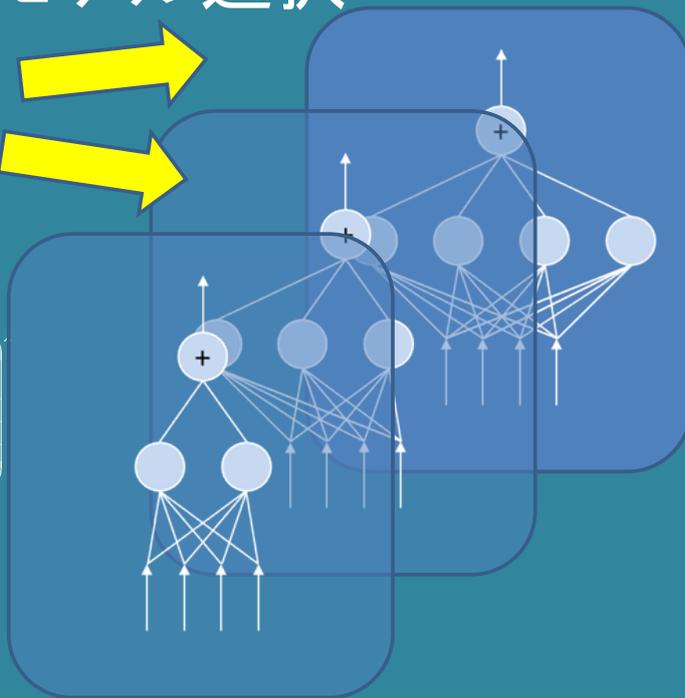
Virtual Concept Drift環境におけるRBFNNのモデル選択

$$E(\mathbf{x}) = \sum_{b=1}^N W(\mathbf{x}) \{F(\mathbf{x}) - f_{\theta}(\mathbf{x})\}^2$$

$$W(\mathbf{x}) = \left(\frac{\hat{q}(\mathbf{x})}{P(\mathbf{x})} \right)^2$$

WLSによる学習

$$= \left[\frac{\Gamma[(N-1+p)/2]}{((N-1)/2)^{p/2} \Gamma[(N-1)/2]} \left[1 + \frac{(\mathbf{x}-\mathbf{u})^T \Sigma^{-1} (\mathbf{x}-\mathbf{u})}{N-1} \right]^{-(N-1+p)/2} \right] \exp\left(-\frac{1}{2} (\mathbf{x}-\mathbf{u})^T \Sigma^{-1} (\mathbf{x}-\mathbf{u})\right)$$



計算量削減方法についても考察を行った。

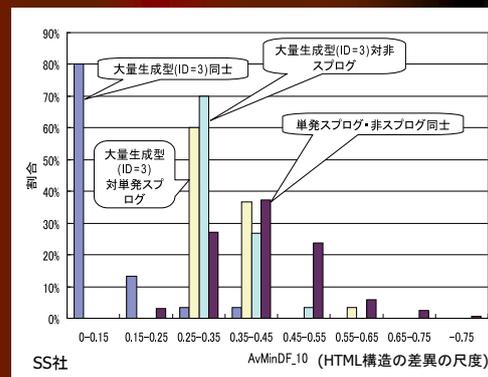
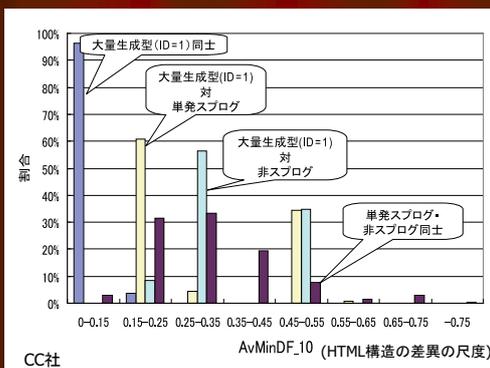
IC_w 最小のモデルを選択

P78 機械学習を用いたスprog検出におけるHTML構造の類似性の利用

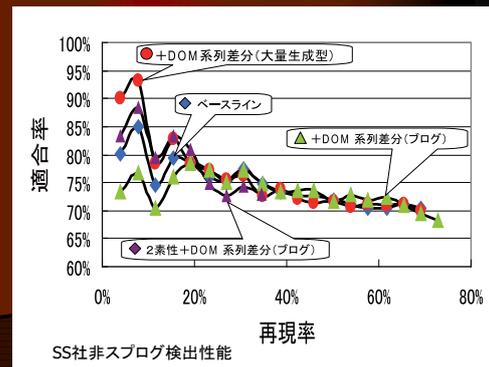
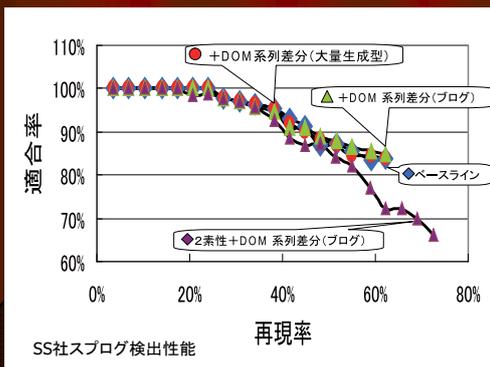
片山 太一, 芳中 隆幸, 宇津呂 武仁,
河田 容英, 福原 知宏

- HTML構造・・・HTML文書から抽出したDOM系列で表現
- 同一スパマーによって大量生成されたスprog・・・DOM系列が類似
- DOM系列の有効性の検証
 - 教師あり学習によるスprog検出
 - 教師なしスprog収集

DOM系列の類似



SVMによるスprog検出



ベータ自由エネルギーとLoopy belief Propagation に現れる グラフのゼータ関数について

渡辺 有祐, 福水 健次

総合研究大学院大学複合科学研究科統計科学専攻
統計数理研究所

Loopy Belief Propagation (LBP) : グラフ構造をもった確率分布の周辺確率分布を計算する近似アルゴリズム。

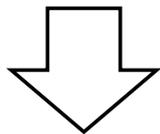
- グラフがツリーだと厳密な値を計算する
- グラフが閉路をもつとき、アルゴリズムの振る舞いは複雑

特に、グラフの形とアルゴリズムの振る舞いの関係を詳しく考える。

主要公式

$$\zeta_G(\mathbf{u})^{-1} = \det(\nabla^2 F) \prod_{ij \in E} \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j) \prod_{i \in V} \prod_{x_i = \pm 1} b_i(x_i)^{1-d_i} 2^{2N+4M}$$

“ベータ自由エネルギー関数Fのヘッセ行列の行列式は、
グラフのゼータ関数 ζ_G の逆数になっている”



この公式からLoopy Belief Propagationの色々な性質が導ける。

- LBPアルゴリズムとベータ自由エネルギーの極小性の詳しい関係
- ベータ自由エネルギー関数の凸性
- 二つの閉路を持つグラフのLBP固定点の一意性について

劣モジュラカットとその応用

No. 89

○河原吉伸(阪大), 永野清人(東工大), 津田宏治(産総研), Jeff Bilmes (Washington Univ.)

「劣モジュラ最大化」

⇒ 多くの機械学習問題を包含

(特徴選択, 能動学習, センサ配置, ...)

L0正則化の場合

$V = \{1, \dots, n\}$ から部分集合 S を選択

$$\max_{S \subseteq V} f(S) \quad \text{s.t. } |S| \leq k$$

劣モジュラ関数

⇒ 集合関数における凸関数に相当

指数関数的に増加する局所解が存在 (計算が極めて困難 (NP-困難))

⇒ 貪欲法に基づく方法が一般的 ((1-1/e)-近似)

厳密解を求める方法論の必要性

- 既存研究は極めて少ない (一般の場合は (Nemhauser & Wolsey, 1981) が唯一)

劣モジュラカットとその応用

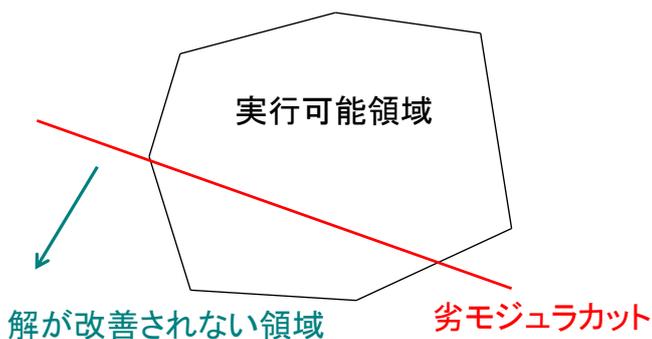
No. 89

○河原吉伸(阪大), 永野清人(東工大), 津田宏治(産総研), Jeff Bilmes (Washington Univ.)

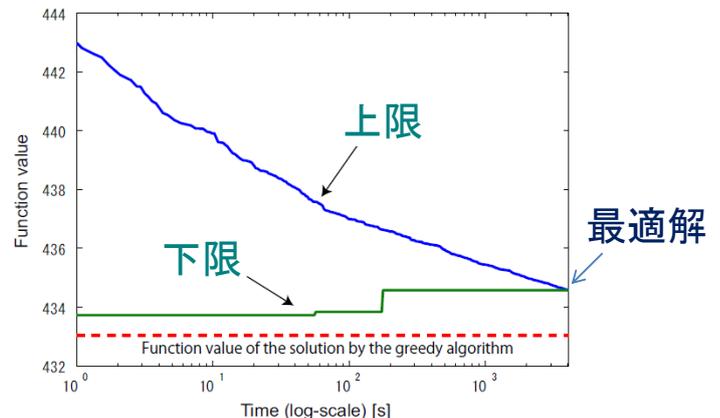
「劣モジュラカット・アルゴリズム」の提案

- 有限回の反復で厳密解へ収束
- 解の上限・下限を与える事が可能 ⇒ ϵ -最適解

カッピング・プレーン法のイメージ



上限・下限の収束の例



(NIPS 2009 においても口頭発表 (スポットライト) として採録)

重み付きカーネルマシンの多次元パス追跡法に関する一考察

烏山 昌幸*, 原田 尚之*, 竹内 一郎*

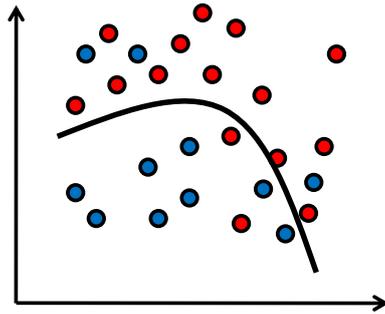
*名古屋工業大学大学院工学研究科

krsym@ics.nitech.ac.jp, harada@goat.ics.nitech.ac.jp, takeuchi.ichiro@nitech.ac.jp

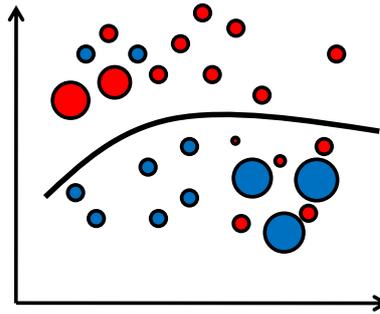
重み付きカーネルマシン: データ点ごとに異なる重みを与えてモデルを推定

データ点ごとの **重要度・信頼度** を反映

(例) 分類問題



重みなし

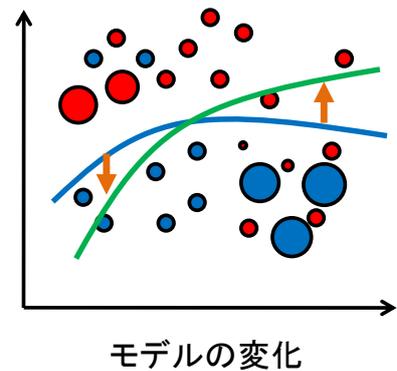
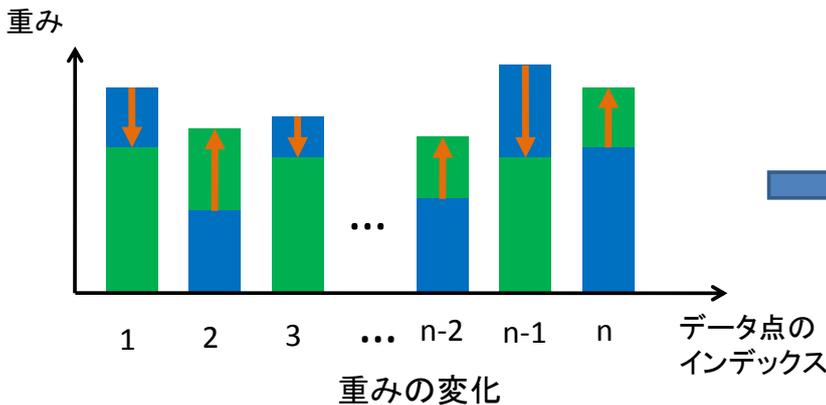


重みあり(大きい点ほど重み大)

与える重みを変化させた時, 推定されるモデルはどう変化するのか?

(重みが変わる事例: 適応的な重み推定, オンライン学習, etc ...)

多次元パス追跡によるモデル更新



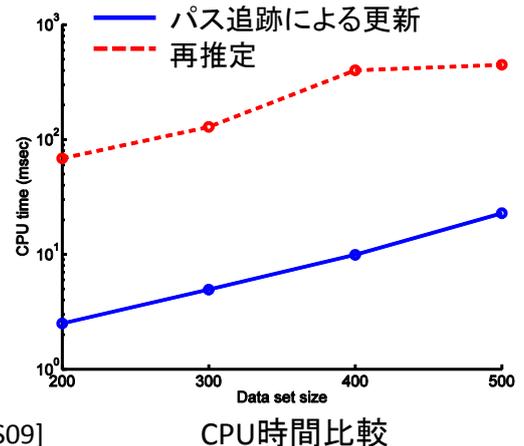
- SVMにおいて重みの変化に対する**モデルの連続的な変化**を解析的に導出(区分線形性)
- 区分線形性を利用したパス追跡による**効率的なモデル更新アルゴリズムの提案**

関連研究:

正則化パス[Hastie et al. JMLR04]

Incremental Decremental Algorithm [Cauwenberghs et al. NIPS01]

Multiple Incremental Decremental Algorithm [Karasuyama, et al. NIPS09]



P099 Split Position Slice Sampler

武井 俊祐,

牧野 貴樹,

高木 利久

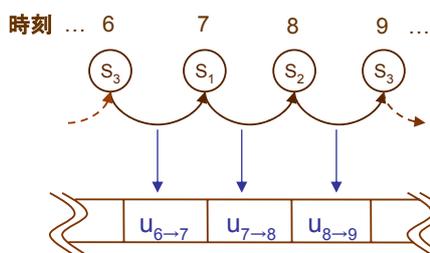
東京大学大学院
新領域創成科学研究科 情報生命科学専攻

東京大学 総括プロジェクト機構

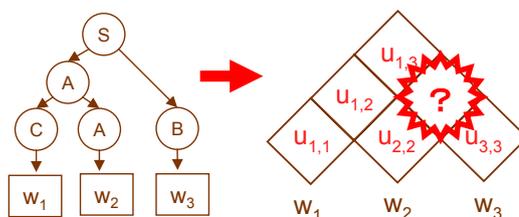
東京大学大学院 新領域創成科学研究科 情報生命科学専攻
東京大学 総括プロジェクト機構
情報・システム研究機構ライフサイエンス統合データベースセンター

- **ベイズPCFGのための高精度かつ高速なMCMC法**
 - 木構造サンプリング(Johnson+, 2007)の高速化
 - **ベイズHMMの高速サンプラー(Beam Sampler; Van Gael+, 2008)を応用**
 - 補助変数の導入による高速な動的計画法
 - **Beam Samplerは直接PCFGには適用できない**
 - 補助変数の導入法を転用できない

Beam Sampler



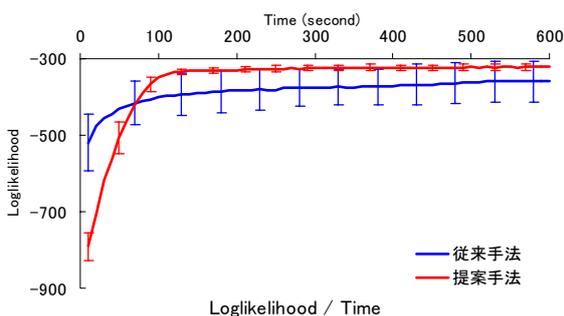
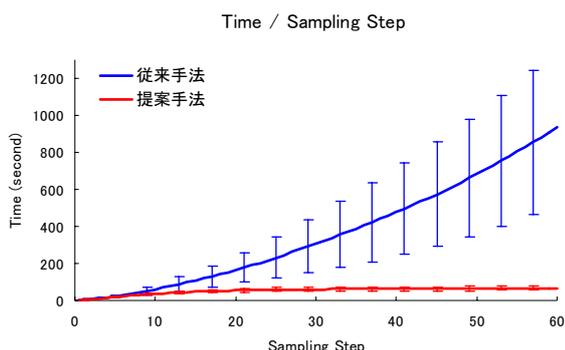
単純な適用



前回のサンプル列から作った補助変数列で動的計画法の各ステップを高速化

動的計画法の全てのステップに補助変数を割り当てられない

→ 枠組みを **Split Position Slice Sampler** へと拡張



上野康隆 赤間陽二

東北大学大学院理学研究科数学専攻

x_1, \dots, x_n : 独立に d 次元標準正規分布に従う

λ_i : 標本分散共分散行列 $S = \sum_{i=1}^n x_i x_i^\top / n$ の第 i 固有値

Theorem (一貫性)

$d = o(n^{\frac{1}{3}})$ のとき, $\sum_{i=1}^k \lambda_i$ ($k \leq d$) は k に確率収束する. すなわち任意の $\varepsilon \geq 0$ に対し $P\left(\left|k - \sum_{i=1}^k \lambda_i\right| \geq \varepsilon\right)$ が $n \rightarrow \infty$ で収束する.

- S は自由度 n , パラメータ E_d/n の Wishart 行列.
- 定理の証明概略. S の固有値の裾確率の評価 (集中不等式) と, 主成分分析が誘導する概念クラスの Vapnik-Chervonenkis 次元の評価.

第12回情報論的学習理論ワークショップ (IBIS 2009)

2009年10月19日(月)-21日(水) 九州大学医学部 百年講堂

ポスターセッションA

ディスカッショントラック



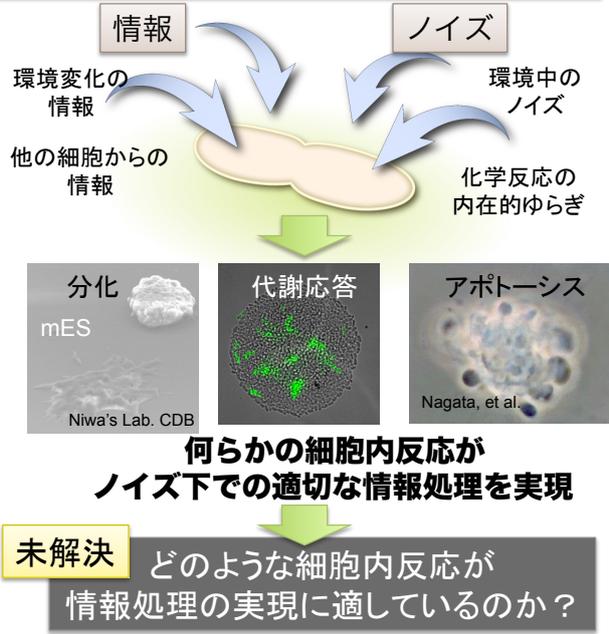
13

ベイズ推定と細胞の運命決定ダイナミクス

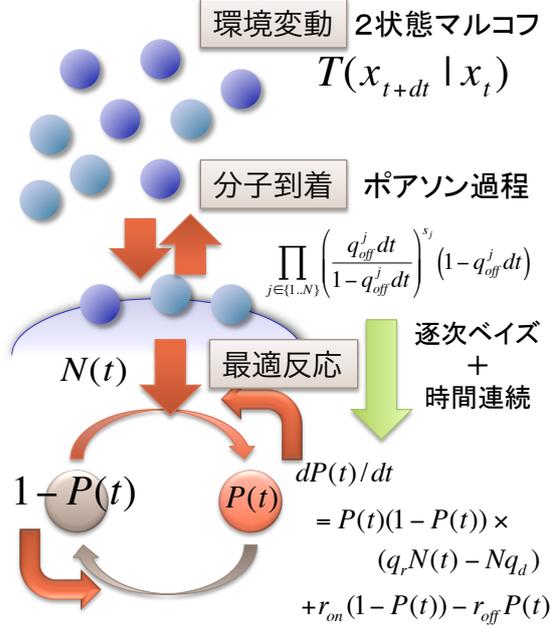
小林徹也：東京大学生産技術研究所 講師：JSTさきがけ研究員



細胞は高ノイズ環境下で運命決定をする

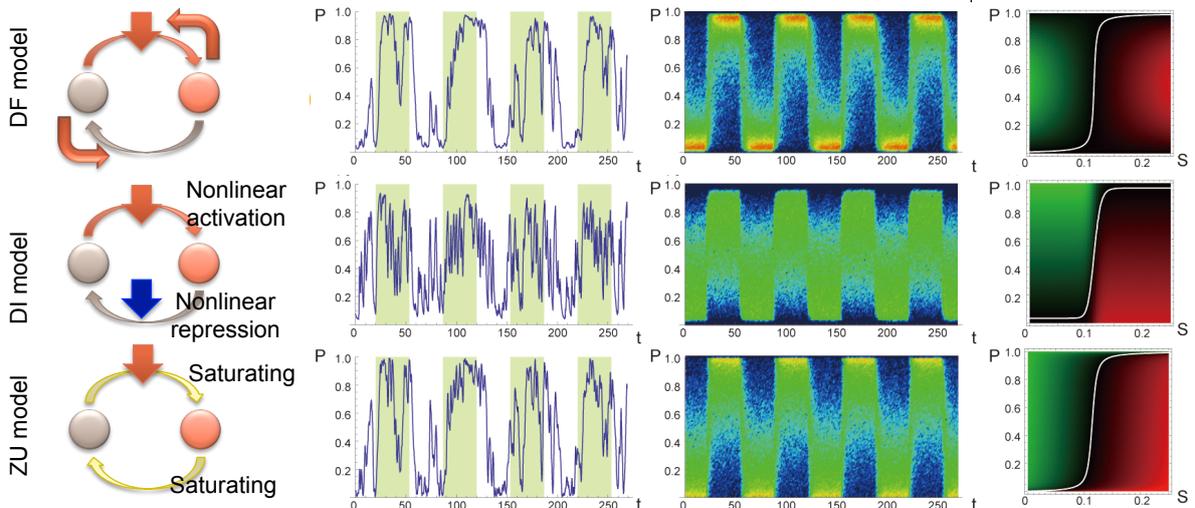


統計推定として最適な反応動態を導く



細胞内反応として実装可能な情報処理動態が明らかになる

各反応ネットワークの環境推定ダイナミクスの比較



最適反応ネットワークと実験知見とが対応



P021 改良型双対尺度法を用いた医薬学データの数量化

山崎 広之¹, 岡本 晃典¹, 日高 伸之介¹, 川下 理日人¹, 高木 達也¹ ttakagi@phs.osaka-u.ac.jp

¹: 大阪大学大学院薬学研究科

Introduction

順序カテゴリーデータの数量化



カテゴリカル解析と比べた場合の、定量的構造活性相関

(QSAR: Quantitative Structure-Activity Relationship) 解析の簡易化及び精度向上



Method

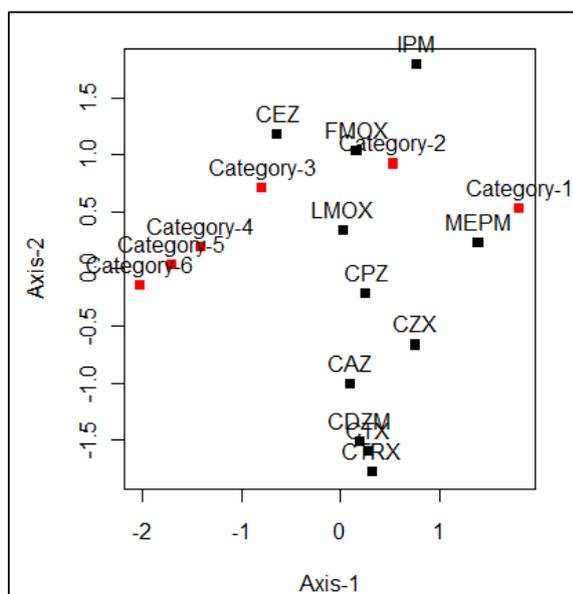
外的基準のない順序カテゴリーデータの数量化の代表例として、双対尺度法が挙げられる[1]。QSAR解析に適切な数量化を行うことを可能とするため、以下の2点の改良を行った。

- 境界値ではなくカテゴリー値を得るようにした。
- 焼きなまし法による、最適なカテゴリー化を可能とした。

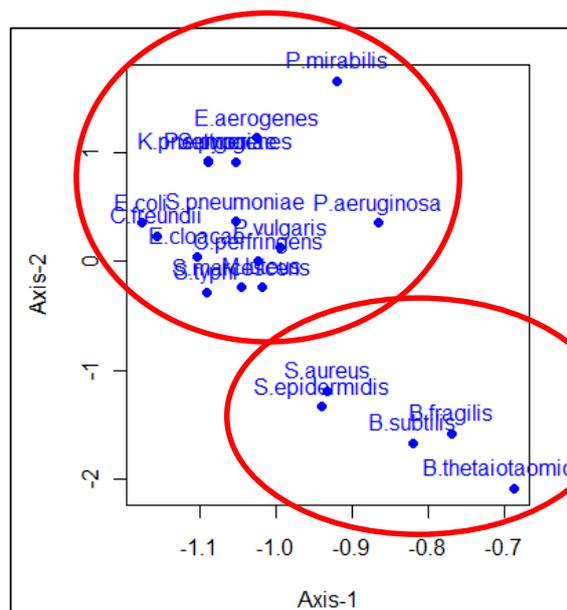
Dataset

医薬品医療機器総合機構添付文書データベース[2]により、抗生物質製剤11種の、計20種の菌種に対する抗菌スペクトルをDatasetとした。

Results



各抗生物質製剤及びカテゴリー値に対する重みの値



菌種に対する重みの値

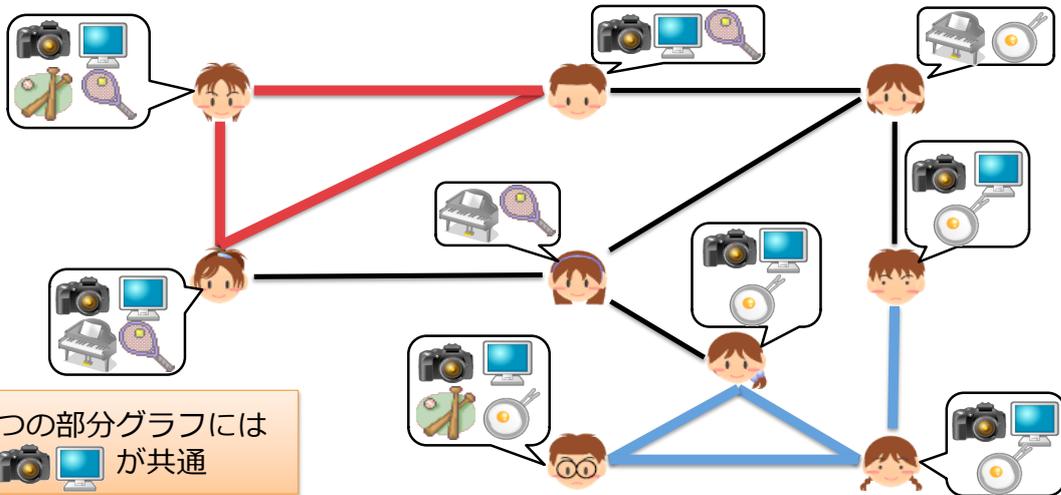
- Reference
- 1.西里静彦著「質的データの数量化 一双対尺度法とその応用」
 2. http://www.info.pmda.go.jp/psearch/html/menu_tenpu_base.html

024 アイテム集合付きグラフからの協調関係抽出

福崎 睦美
お茶の水女子大学

鹿島 久嗣
東京大学

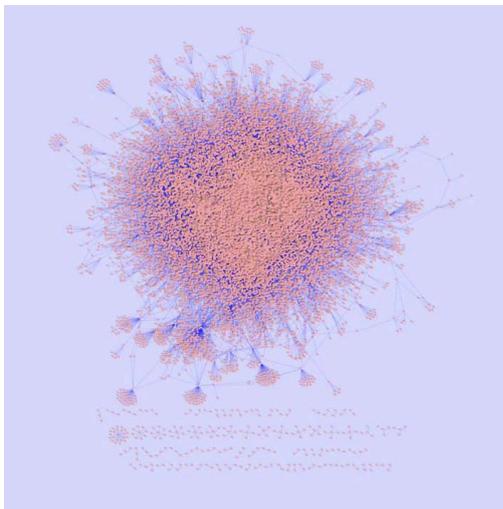
瀬々 潤
お茶の水女子大学



- 本来は一つのグループであり、グループ間には未発見の関係がある
または、
 - 一方のグループによく購入される商品ももう一方のグループにも購入されやすい
- 共著者ネットワークを解析→同じ著者が研究しているプロジェクトを発見
遺伝子ネットワークを解析→薬剤副作用発見などにも応用可能

共著者ネットワークの解析結果

ノード:論文 辺:リファレンス
アイテム:著者



結果の一例

抽出されたそれぞれの部分グラフに
含まれる論文の研究分野

ビデオ-オン-デマンド
ストレージシステム

マルチ
データベース

メインメモリ
データベース

著者
Rajeev Rastogi
Abraham Silberschatz

P032: ベイズ推定を用いない曲指数型分布族の推定量の改善

九州大学 川喜田 雅則, 竹内 純一

背景 plug-in 分布は m 埋め込み曲率方向へ曲率に応じてシフトすることで期待 KL を最適に改良できること、またベイズ予測分布はこのシフトを事前分布に依らずに達成することが知られている (Komaki, 1996)。しかし高次元の場合はモデルに特別な構造を仮定しない限りベイズ予測分布の構築は困難。

新しい予測分布の提案

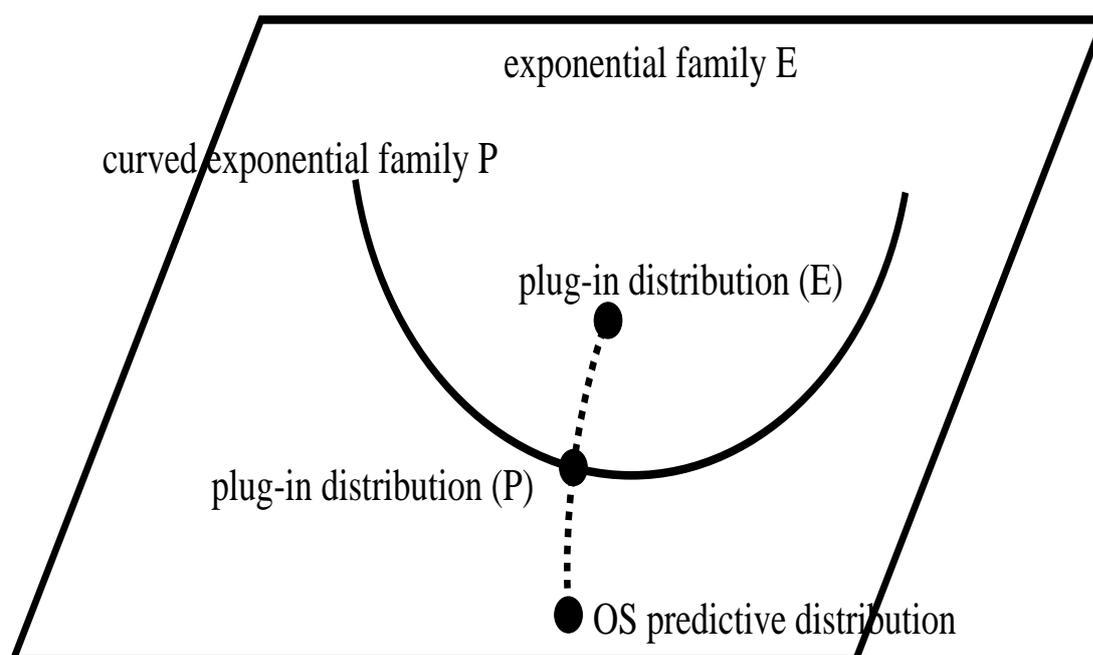
- η 座標系において曲指数型分布族のプラグイン分布から外側の指数型分布族のプラグイン分布へのシフトと正反対にシフトする **Oppositely-Shifted 予測分布** を提案
- OS 予測分布は指数型分布族から飛び出せないという制約の下での最適なシフトを平均的に達成することを示す。

October 4, 2009

ベイズ推定を用いない曲指数型分布族の推定量の改善

1

OS 予測分布の構成法



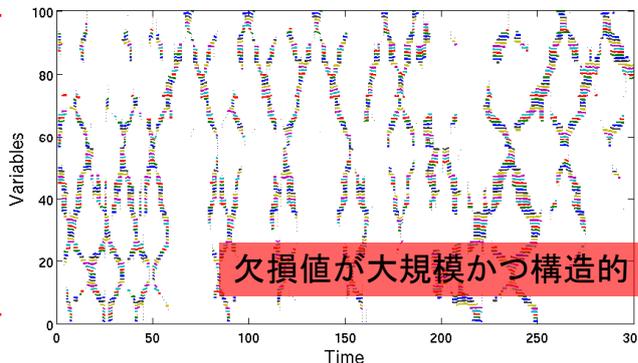
October 4, 2009

ベイズ推定を用いない曲指数型分布族の推定量の改善

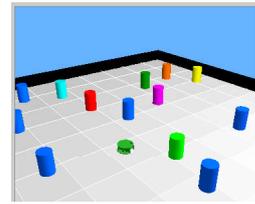
2

(P038) 構造的欠損値を伴う多次元時系列データのための制約付き共埋め込み法
矢入健久(東京大学)

この研究で扱うデータ

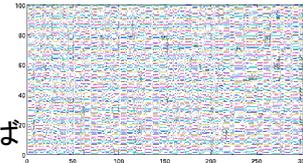


生成



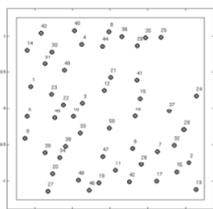
移動ロボットによる物体観測
(相対方位角)の履歴

(参考)もし
欠損なければ



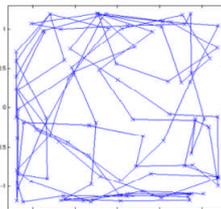
やりたいこと: 物体地図とロボットの状態軌跡の同時復元(SLAM問題)

制約
付き
次元
削減



地図(物体の位置)

制約
付き
次元
削減



軌跡(ロボットの状態遷移)

同時次元削減
(共埋め込み)問題

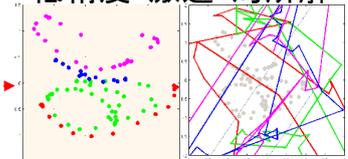
完全データなら容易
だが、構造的欠損
データに対しては?

(P038) 構造的欠損値を伴う多次元時系列データのための制約付き共埋め込み法
矢入健久(東京大学)

(代表的)従来手法: PCAMD (Wiberg algorithm or ALS)

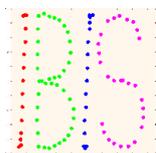
- 計算機視覚におけるStructure from Motion (SfM)
- 不完全データに対する特異値分解 (SVD)
- ランダムな欠損を前提。反復法

低精度・激遅・局所解

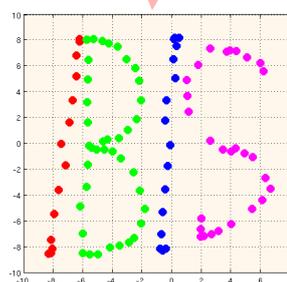


提案手法: 局所線形アラインメント法(仮)

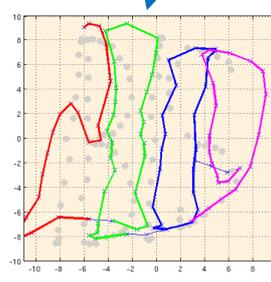
- Weighted Multiple CCA (?)
- 欠損の構造的パターンを利用
- 様々な制約の利用が容易
- 固有値あるいは線形方程式一発



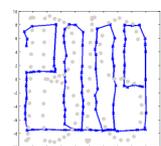
正解



推定された地図



推定された軌跡



正解

P047 混合Bernoulli分布に基づく変分Bayes法による連想記憶モデルの解析

荒木佑季, 永田賢二, 岡田真人, 井上真郷

早稲田大学 東京大学 東京大学 早稲田大学

ボルツマン分布

$$p(\mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{1}{T} H(\mathbf{x})\right)$$

• Hebb則

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (\xi \in \{+1, -1\}^N)$$

• ハミルトニアン

$$H(\mathbf{x}) = -\frac{1}{2} \sum_i x_i \sum_j J_{ij} x_j$$

MCMC法
+1→0
-1→1
に交換

高次元データ

$$\mathbf{x}^0 = (x_0^0, x_1^0, \dots, x_N^0)$$

$$\mathbf{x}^1 = (x_0^1, x_1^1, \dots, x_N^1)$$

⋮

$$\mathbf{x}^M = (x_0^M, x_1^M, \dots, x_N^M)$$

混合ベルヌーイ分布による変分Bayes法

クラスタリングにより系のクラスタ数を推定

- Bayes予測分布を近似的に求める
- 事後確率最大化(MAP)推定を行うことでモデル選択

* クラスタ数は次元数(ニューロンの数)N, 記憶パターン数p, 温度パラメータTの条件で変化

• 確率モデル

$$p(\mathbf{x}, z | a, b, K) = \prod_{m=1}^M \prod_{k=1}^K \left(a_k \prod_{n=1}^N b_{k,n}^{x_{m,n}} (1-b_{k,n})^{1-x_{m,n}} \right)^{z_{m,k}}$$

• 混合割合の共役事前分布: Dirichlet分布

$$p(a | K) = \delta\left(\sum_{k=1}^K a_k - 1\right) \frac{\Gamma\left(\sum_{k=1}^K \tilde{\phi}_k\right)}{\prod_{k=1}^K \Gamma(\tilde{\phi}_k)} \prod_{k=1}^K a_k^{\tilde{\phi}_k - 1}$$

• Bernoulli分布の共役事前分布: ベータ分布

$$p(b | K) = \prod_{k=1}^K \prod_{n=1}^N \frac{\Gamma(\tilde{\beta}_{0,k,n} + \tilde{\beta}_{1,k,n})}{\Gamma(\tilde{\beta}_{0,k,n}) \Gamma(\tilde{\beta}_{1,k,n})} b_{k,n}^{\tilde{\beta}_{0,k,n} - 1} (1-b_{k,n})^{\tilde{\beta}_{1,k,n} - 1}$$

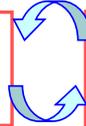
VBアルゴリズム

<VB-Eステップ>

• 潜在変数更新
 $z_{m,k} : m = 1, \dots, M, k = 1, \dots, K$

<VB-Mステップ>

• パラメータ更新
 $\phi_k : k = 1, \dots, K$
 $\beta_{0,k,n}, \beta_{1,k,n} : k = 1, \dots, K, n = 1, \dots, N$



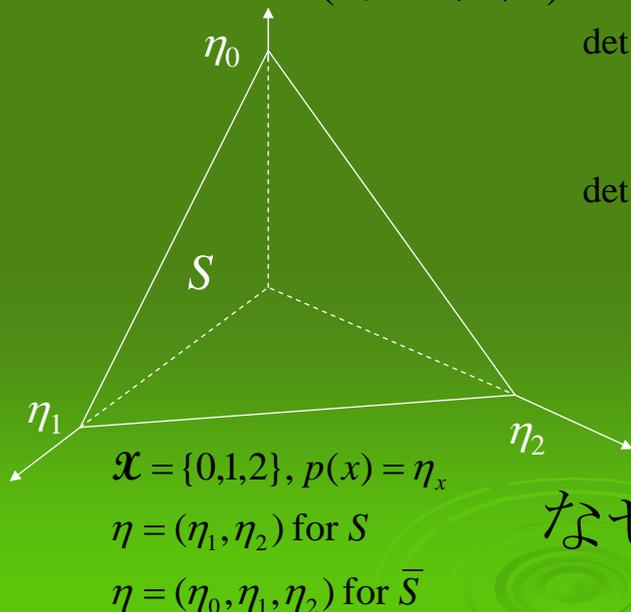
P61 定常Markovモデルの体積要素と 拡大モデルについて 竹内純一



Math-for-industry
Education & Research Hub

多項ベルヌーイモデルの場合

\bar{S} : 拡大モデル(非正規化)



$$\det J = \begin{vmatrix} 1/\eta_1 + 1/\eta_0 & 1/\eta_0 \\ 1/\eta_0 & 1/\eta_2 + 1/\eta_0 \end{vmatrix} = \frac{1}{\eta_0 \eta_1 \eta_2}$$

$$\det \bar{J} = \begin{vmatrix} 1/\eta_0 & 0 & 0 \\ 0 & 1/\eta_1 & 0 \\ 0 & 0 & 1/\eta_2 \end{vmatrix} = \frac{1}{\eta_0 \eta_1 \eta_2}$$

J : Fisher情報行列

S の体積要素(密度) $\sqrt{(\det J)}$

S のJeffreys事前分布 $w(\eta) \propto \sqrt{(\det J)}$

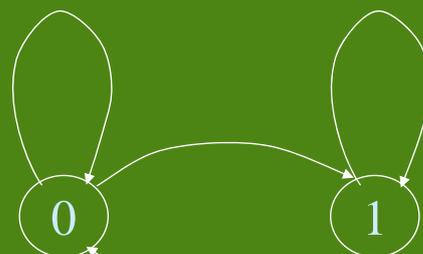
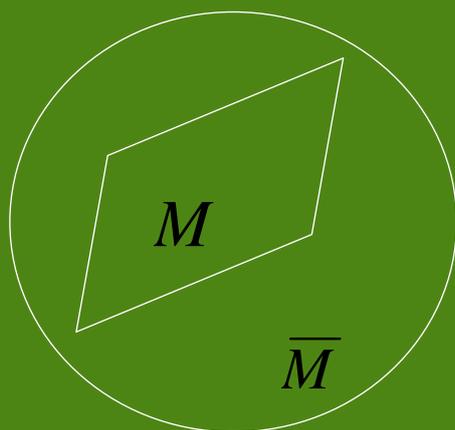
なぜ $\det J = \det \bar{J}$ か?

マルコフ連鎖モデルの場合



Math-for-industry
Education & Research Hub

2状態の例



$\mathcal{X} = \{0,1\}$

$\eta = (\eta_{00}, \eta_{10})$

$\bar{\eta} = (\eta_{00}, \eta_{01}, \eta_{10}, \eta_{11})$

$$\det J = \left(\sum_{x \in \mathcal{X}} \Delta_{xx} \right)^2 \prod_{x \in \mathcal{X}} \frac{\eta_x}{\prod_{y \in \mathcal{X}} \eta_{xy}}$$

$$\det \bar{J} = 0$$

[Takeuchi 2009]

拡大モデルでは
 \bar{J} が退化する

リスク考慮型強化学習に向けたリターン分布推定

森村 哲郎[†], 杉山 将^{††}, 鹿島 久嗣[‡], 八谷 大岳^{††}, 田中 利幸^{††}

[†]: IBM東京基礎研究所, ^{††}: 東京工業大学, [‡]: 東京大学, ^{††}: 京都大学

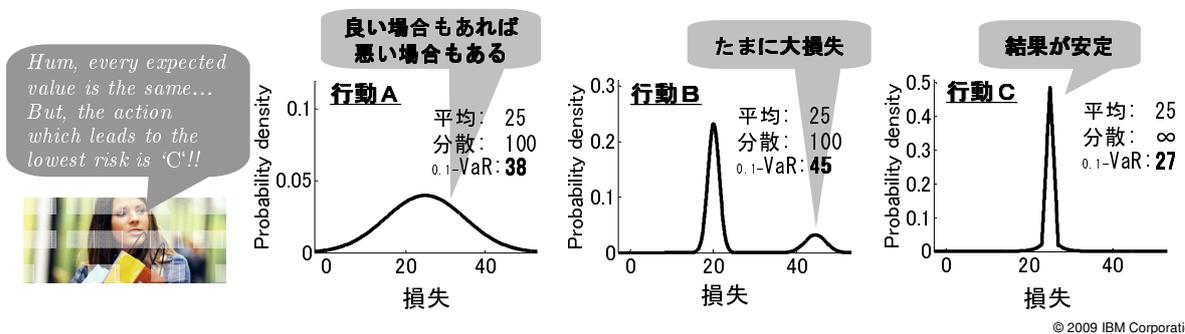
*tetsuro@jp.ibm.com

背景

- 意思決定は、各選択肢のもたらす「損失」の期待値をもとに行われるのが普通だが、思いがけず起こる大損失のリスクを回避するには、損失の分布を考える必要がある
- これを逐次的な意思決定の枠組みの中で扱うには、累積損失の分布推定問題を解くことが必要。しかし、単純な密度推定法では効率が悪く現実的ではない

成果：リスク回避型意思決定の効率的な実現手法の提案

- 強化学習におけるTD学習の様式に基づく、効率よい累積損失分布の推定法の提案



アプローチ

- リターン分布 $p_\eta(\eta|s)$ におけるBellman方程式(再帰的な式)を利用 [田中&中田, 2006]

$$p_\eta(\eta|s) = \mathbf{T} p_\eta(\eta|s)$$

$$\left(\begin{array}{l} \mathbf{T}: \text{分布ベルマンオペレータ} \\ \mathbf{T} p_\eta(\eta|s) \triangleq \frac{1}{\gamma} \sum_{s', a} p_\Gamma(s'|s, a) p_a(a|s) \int_r p_r(r|s, a, s') p_\eta(\eta' = \frac{\eta - r}{\gamma} | s') dr \end{array} \right)$$

- 近似リターン分布 $\hat{p}_\eta(\eta|s)$ を $\mathbf{T} \hat{p}_\eta(\eta|s)$ に近づけることで学習
- パラメトリック推定: パラメータ θ でリターン分布を表現 $\hat{p}_\eta(\eta|s, \theta)$
 - $\mathbf{T} \hat{p}_\eta(\eta|s, \theta)$ から $\hat{p}_\eta(\eta|s, \theta)$ のKullback-Leiblerダイバージェンスを小さくする。(自然勾配法を用いて、パラメータ θ を最適化)
- ノンパラメトリック推定: パーティクルでリターン分布を表現
 - 下記の漸近性に基づき、現時刻のパーティクルを一時刻先のパーティクルで更新する。(パーティクル・スムージングにより、パーティクルを最適化)

$$p_\eta(\eta|s) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \delta(r_n + \gamma \eta'_n - \eta).$$

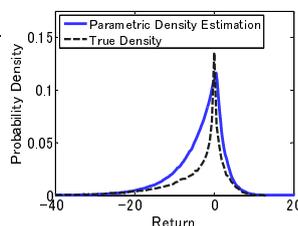
実験結果: 6状態, 4行動 MDP

—: 推定密度分布, - - -: 真のリターン密度分布

✓ パラメトリック推定

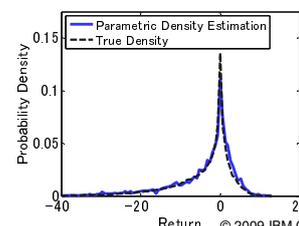
(非対称ラプラス分布を使用)

収束に必要な試行数: **少**
モデル自由度: **低**



✓ ノンパラメトリック推定

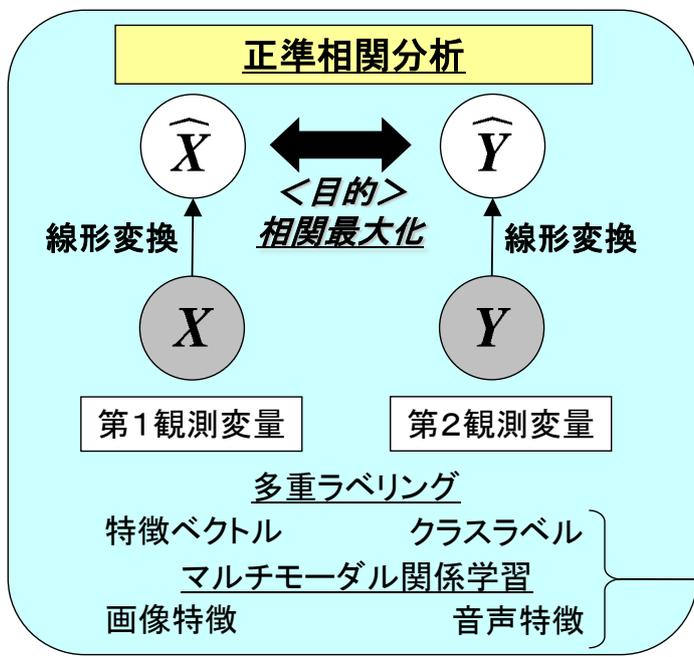
収束に必要な試行数: **多**
モデル自由度: **高**



SemiCCA: Efficient semi-supervised learning of canonical correlations

木村昭悟⁽¹⁾、亀岡弘和⁽¹⁾、杉山将⁽²⁾、前田英作⁽¹⁾、坂野鋭⁽¹⁾、石黒勝彦⁽¹⁾

(1) 日本電信電話(株) NTTコミュニケーション科学基礎研究所 (2) 東京工業大学 情報理工学研究科



提案する枠組: SemiCCA

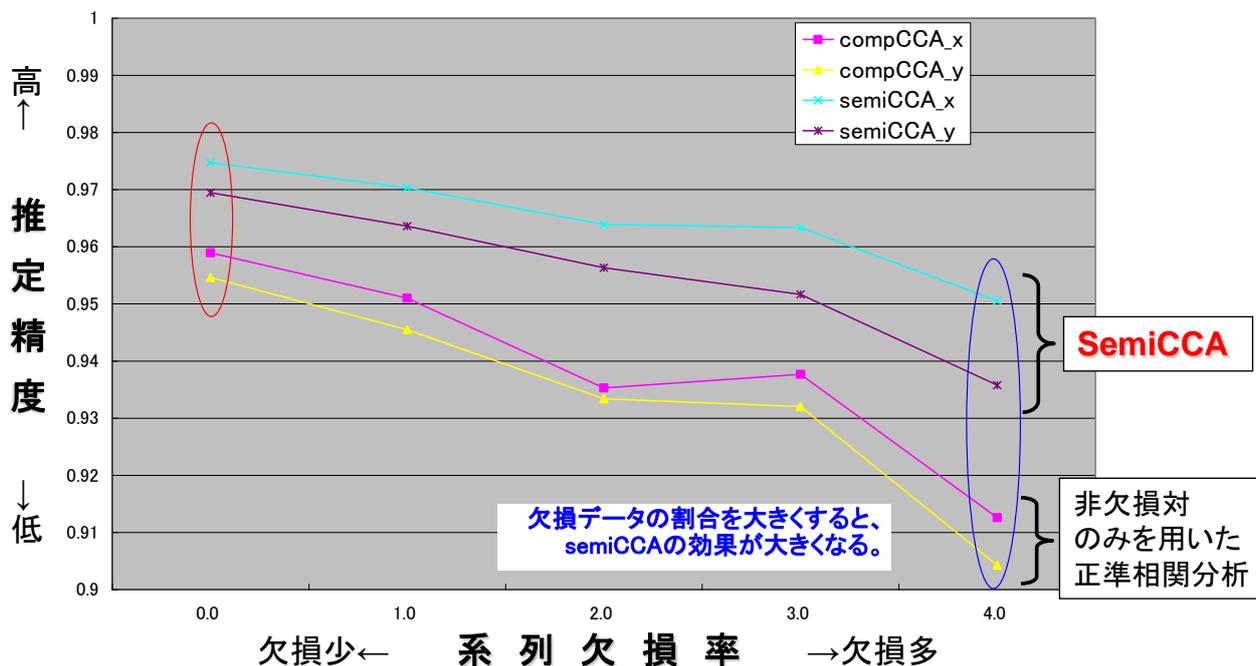
- **正準相関分析を“半教師化”**
共起する多次元ベクトル対の集合において、対の一方のベクトルが大量に欠損していたとしても、それらの相関関係を高精度で推定
- **簡単に計算ができる**
正準相関分析とほぼ同様の計算コストで計算が可能

必ずしも両方が揃っているとは限らない

実験結果

詳細は [P84] にてご紹介致します。

- 10000個の多次元系列対を人工的に生成、一方の系列を人為的に欠損させ、欠損率と推定精度との関係を調査



No.87

飽和性と極大性を用いた単一系列データにおける 頻出系列の高速マイニング

村田拓也

岩沼宏治

山梨大学医学工学総合教育部

山梨大学医学工学総合教育部

研究目的:

単一系列データを効率的にマイニング
抽出される大量の頻出系列を圧縮

飽和性

- 同じ出現頻度を持つ他の頻出系列の部分系列とならない系列のみを抽出
- すべての頻出系列とその頻度を復元できる

極大性

- 他の頻出系列の部分系列とならない系列のみを抽出
- 頻出系列の頻度情報は復元できない

実験結果(抽出系列数)

DB①

- Webアクセスログデータ
- 系列長: 5000
- アイテム系列

最小サポート	頻出系列	頻出右飽和系列	頻出右極大系列
0.04%	75,242	10,731	5,589
0.06%	11,985	5,550	3,347
0.08%	4,563	3,208	2,072
0.1%	2,472	2078	1,375

DB②

- 毎日新聞記事データ1年分
- 系列長: 365
- アイテム集合系列

最小サポート	頻出系列	頻出右飽和系列	頻出右極大系列
5%	162,131	162,002	138,161
6%	77,015	76,969	66,088
7%	32,928	32,917	28,459
8%	15,754	15,751	13,648

データベースの特徴により、あまり圧縮できない場合もある

参考: Wang, J, and Han, J, BIDE: Efficient Mining of Frequent Closed Sequences. *Proc. Int. Conf. on Data Engineering (ICDE'04)*, pp.79-90 (2004)

多次元尺度法を用いたアドレナリン受容体の機能分類

井上 泰仁

舞鶴工業高等専門学校 電気情報工学科

研究目的

- アドレナリン受容体は、カテコールアミン類によって活性化される受容体である。
- リガンドの種類、Gタンパク質の種類から、 $\alpha 1$ 、 $\alpha 2$ 、 β の3種類に分類される。
- 結合するGタンパク質が未知なタンパク質の解明に役立つ。

使用したデータと解析方法

タンパク質データベースUni-Prot (Release 56)



アドレナリン受容体入手 (57配列)

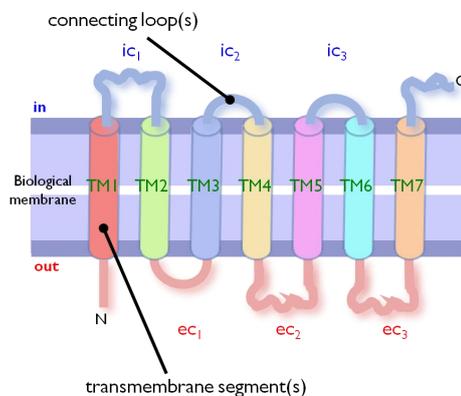
ループ領域のアミノ酸残基数を**多次元尺度法**で解析

Figure 1 Nomenclature for loops and N-/C-terminal loops of adrenergic receptors (a) Crystal structure of Rhodopsin. (b) Numbers are assigned from N- to C-terminal to each extracellular loop (ec_1 – ec_3), intracellular loop (ic_1 – ic_3), and transmembrane region (TM1–TM7). Transmembrane regions are indicated as syringes in the biological membrane.

解析結果

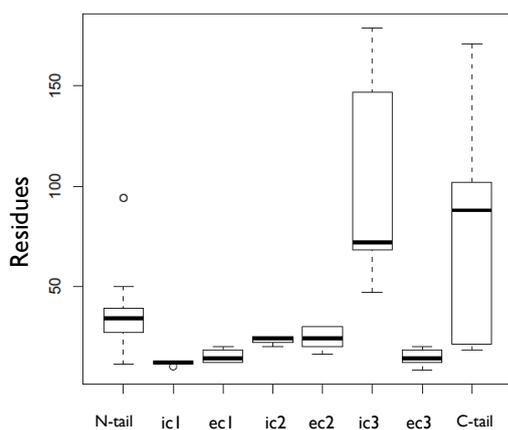


Figure 2 Boxplot for the lengths of N-/C-tail, intracellular (ic_1 – ic_3), and extracellular (ec_1 – ec_3) loops.

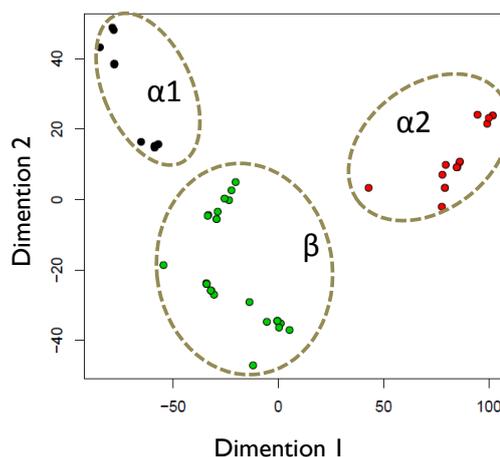


Figure 3 Scatter plot of adrenergic receptors using the MDS method. Black: adrenergic $\alpha 1$ receptors, Red: adrenergic $\alpha 2$ receptors, Green: adrenergic β receptors.

参考文献

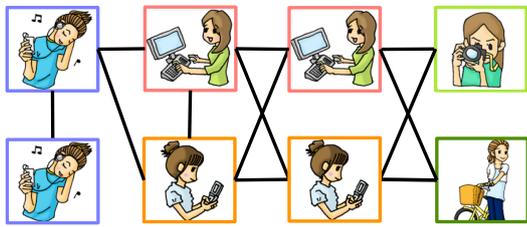
Inoue, Y., Ikeda, M., and Shimizu, T., Proteome-wide classification and identification of mammalian-type GPCRs by binary topology pattern, *Computational Biology and Chemistry*, 28(1), 39-49, 2004.

ポスター番号：94

つながり方に着目した属性付きグラフの概要構造抽出

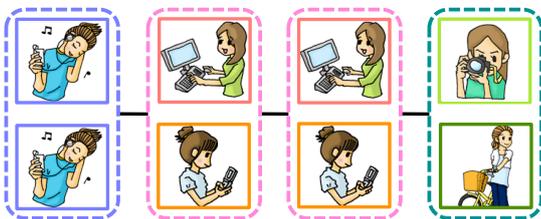
寺田愛花（お茶の水女子大学），瀬々潤（お茶の水女子大学）

例：SNS



	生物ネットワーク	Web	SNS
頂点	タンパク質, 遺伝子	Webページ	ユーザ
属性	機能	カテゴリ	コミュニティ
辺	相互作用	リンク	友人関係

など…



- グループ内の属性は少ない
- グループ間に辺がある
→ できるだけ多くの頂点が相互作用
- 同一の属性でも他との関係が異なるものは別のグループ
- 異なる属性でも他との関係が類似しているものは同一のグループ

最適化問題の提案と解析結果

指標の定義

\mathcal{C} ：グラフのグループ集合、 C_i ：頂点のグループ

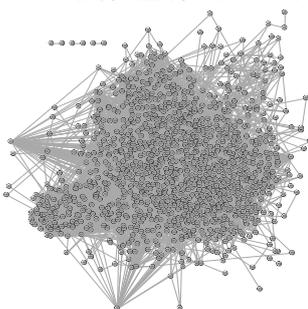
$$w(\mathcal{C}) = \sum_{C_i \in \mathcal{C}} \left(Ent(C_i) \times \sum_{C_j \in \mathcal{C} \text{ and } C_j \neq C_i} c(C_i, C_j) \right) - \lambda |\mathcal{C}|^2$$

最大化

$Ent(C_i)$ ：グループ C_i の属性が同一であるほど大きい値

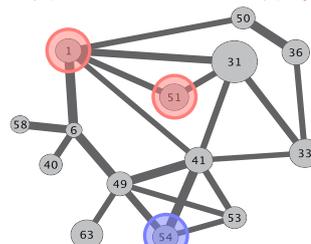
$c(C_i, C_j)$ ：グループ C_i と C_j の関係があるほど大きい値

タンパク質相互作用ネットワーク



頂点：タンパク質（1,184個）
属性：タンパク質の機能（221種類）
辺：相互作用（9,807本）

同じ機能でも役割によって
二つ以上のグループに分類



異なる機能でも役割が類似したものを
同一グループに分類

頂点：タンパク質のグループ
辺：関係が強いグループ

P97 教師なしアンサンブル適応法の提案と音響モデル適応への応用

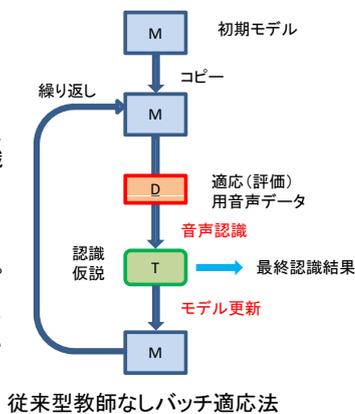
篠崎 隆宏, 久保田 雄, 古井 貞熙

東京工業大学 計算工学専攻 www.furui.cs.titech.ac.jp

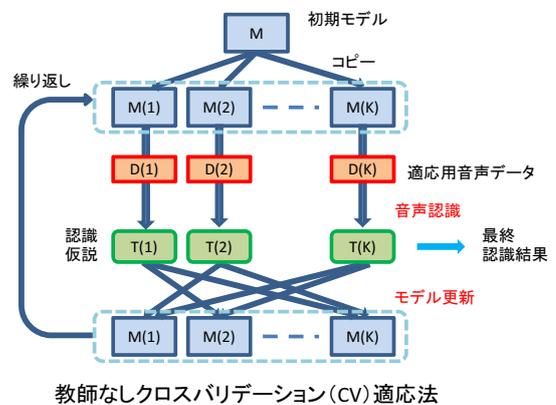


- 音声認識において、**バッチ型教師なし適応**は人手による書き起こしデータを用いずに高い認識精度を得る手法として広く用いられている。
- バッチ型教師なし適応では音声認識により認識仮説を求め、それをラベルとして用いてモデル更新を行うことを繰り返す。**認識仮説には誤りが含まれるため、適応化性能が限られる**問題がある。
- アンサンブル的手法を教師なし適応に応用することで適応化性能を向上させることを目的とし、**教師なしクロスバリデーション(CV)適応法**および**教師なし集合(Ag)適応法**を提案した。大語彙音声認識実験により顕著な適応性能改善効果があることを示す。
- 提案法は一般性があり、音声認識に限らず**幅広い応用**が考えられる。

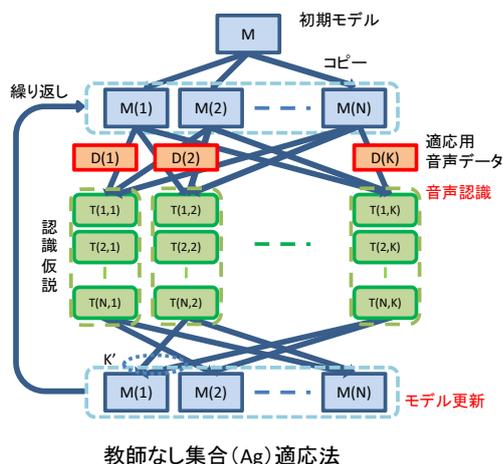
従来型バッチ型適応では、初期モデルを用いた認識対象音声の認識と、認識結果を用いたモデル更新を繰り返す。モデル更新には種々の手法が用いられる。



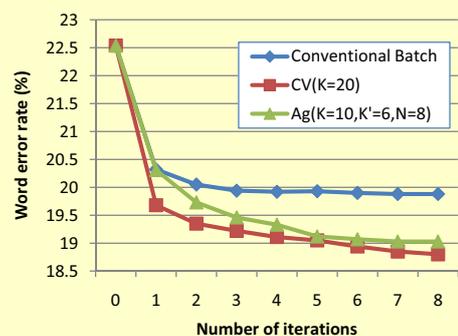
教師なしCV適応では、モデル適応に用いる認識対象音声データをK個の排他的なサブセットに分割する。認識ステップとモデル更新ステップで使用するデータをCV法により分離することで繰り返しループ中での認識誤りの影響を低減し、適応化性能の向上を図る。



教師なし集合適応では、K個に分割した認識対象音声のサブセットそれぞれをN個のモデルを用いて認識する。K個中ランダムに選んだK'個のサブセットに対応するNK'個の認識結果を用いてN個のモデルそれぞれを更新する。N個のモデルを用いた認識結果を同時にモデル更新に用いることで、個々のモデルによる認識誤りの影響を低減し、適応化性能の向上を図る。



実験結果 (音響モデルの教師なし話者適応)



関連文献

• T. Shinozaki, Y. Kubota, and S. Furui, "Unsupervised cross-validation adaptation algorithms for improved adaptation performance," ICASSP, pp. 4377-4380 (2009).
 • 篠崎 隆宏, 「音響モデルのアンサンブル学習」, 2009-SLP-77, No. 11 (2009-7).

第12回情報論的学習理論ワークショップ (IBIS 2009)

2009年10月19日(月)-21日(水) 九州大学医学部 百年講堂

ポスターセッションB



第12回情報論的学習理論ワークショップ (IBIS 2009)

2009年10月19日(月)-21日(水) 九州大学医学部 百年講堂

ポスターセッションB

テクニカルトラック



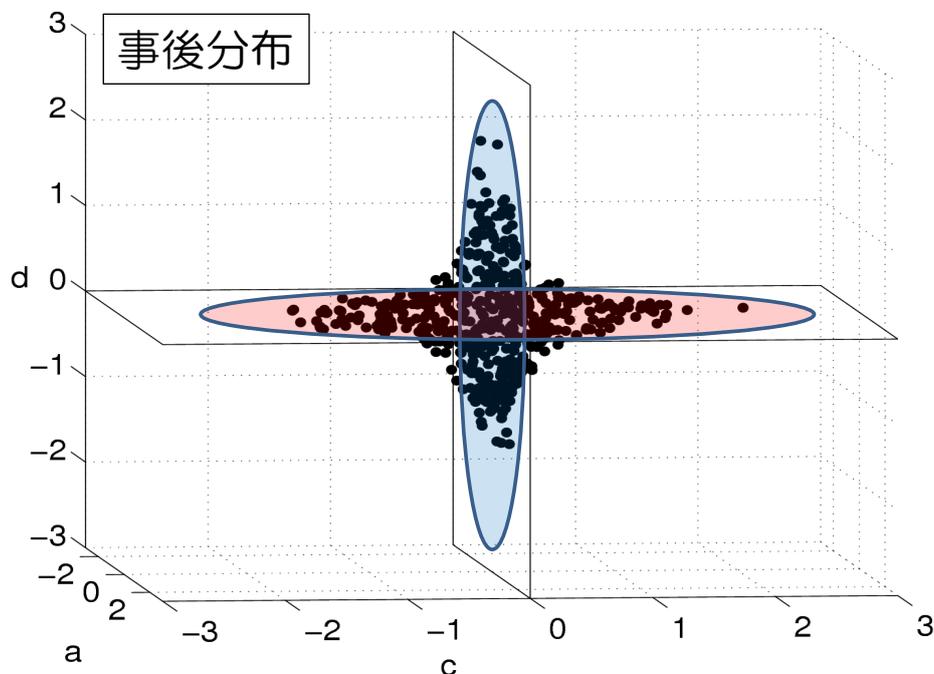
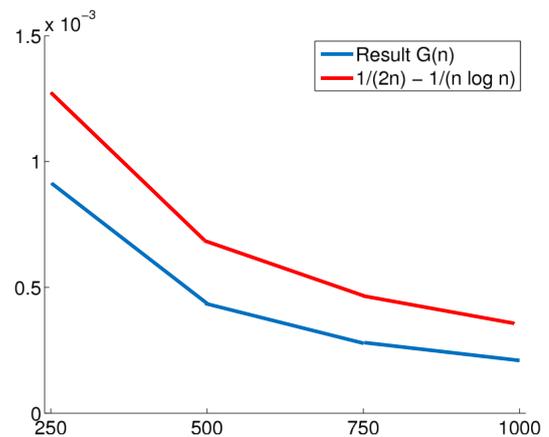
P004

1次元線形動的システムの特異性と バイズ汎化誤差への影響

内藤 卓人, 山崎啓介 (東工大・渡辺研)

線形動的システムの**バイズ汎化誤差**を
パラメータ空間の特異性を利用して**説明**.

$$G(n) \simeq \frac{\lambda}{n} - \frac{m-1}{\log \log n}$$
$$\leq \frac{1}{2n} - \frac{1}{\log \log n}$$



真のパラメータが
二つの平面上に分布



カルマンフィルタによる
状態推定が一意に定まらない!

P010

品質の異なる二つのデータ集合間の の転移学習の解析

An Analysis of Transfer Learning between
a Pair of Datasets with Different Qualities

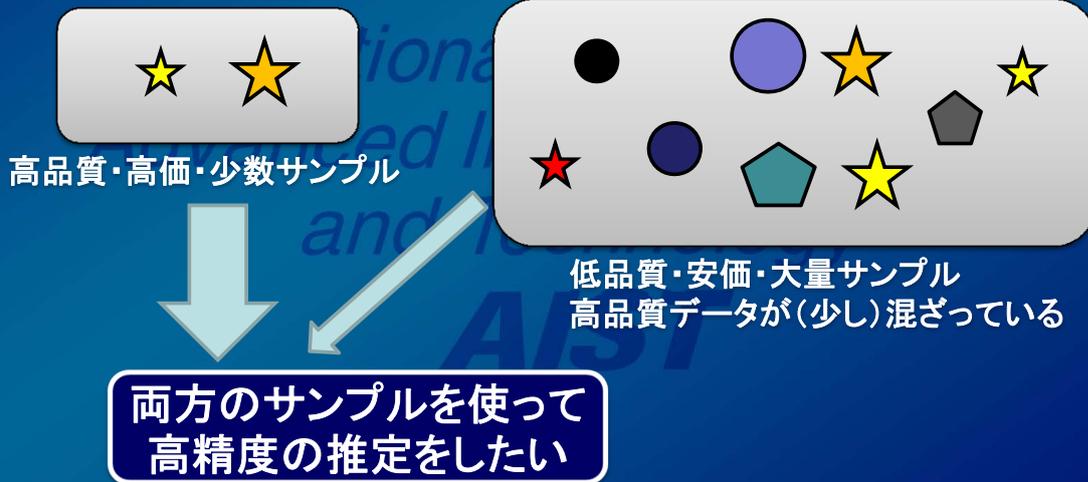
赤穂昭太郎

神島敏弘

Shotaro Akaho

Toshihiro Kamishima

産総研 s.akaho@aist.go.jp, mail@kamishima.net

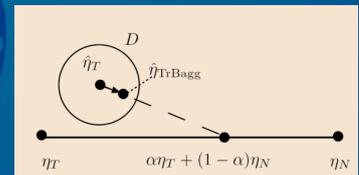


単純なモデルでの解析

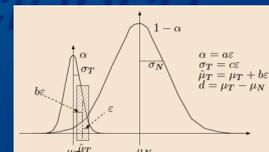
- バギングに基づいたアルゴリズム(神島)
- タグ付け問題で良好な結果(Kamishima et al. ICDM2009など)

どんなときにうまくいく・いかないのか？

- **結果1:**
単純な指数分布族の推定問題ではうまくいかない
(ゴミに引っ張られる)



- **改良アルゴリズムの提案**
- **結果2:**
正規分布の位置パラメータの推定
改良アルゴリズムが緩い条件で真の値との誤差を小さくできる



$$\gamma \simeq 2 \frac{\phi(d, \sigma_N^2)}{a} + 2\phi(1, c^2)$$

P11: 条件付きエントロピー最小化に基づく 教師付き次元削減手法

日野英逸, 村田昇
早稲田大学

- 教師付き次元削減のための基準として, 条件付きエントロピー最小化を用いることを提案.
- カーネル密度推定を介してエントロピーを推定・最小化することで, 可視化, 判別に有効な低次元データ表現を得る.

- カーネル法により非線型次元削減に拡張.
- 条件付きエントロピー最小化基準を, multiple kernel learning(MKL)問題にも適用.
 - 新しいMKLの枠組み.
 - 実データを用いた判別実験の結果, 既存の非線型次元削減(KFDA)より優れた結果が得られた. 既存のMKL手法との比較では, state-of-the-artな方法と同程度の精度が得られた.

P033 : ネットワーク科学の方法を用いたWebページ ネットワークの構造による分類

中川 帝人、鈴木 泰博(名古屋大学情報科学研究科) nakagawa.teito@b.mbox.nagyoa-u.ac.jp

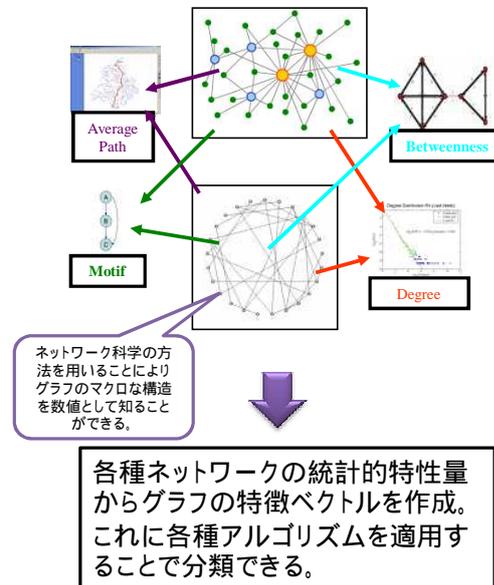
問題: グラフ分類問題

目的: 複数のグラフをマクロな構造から分類したい。

アイデア: ネットワーク科学の統計的特性量をグラフの特徴とする。

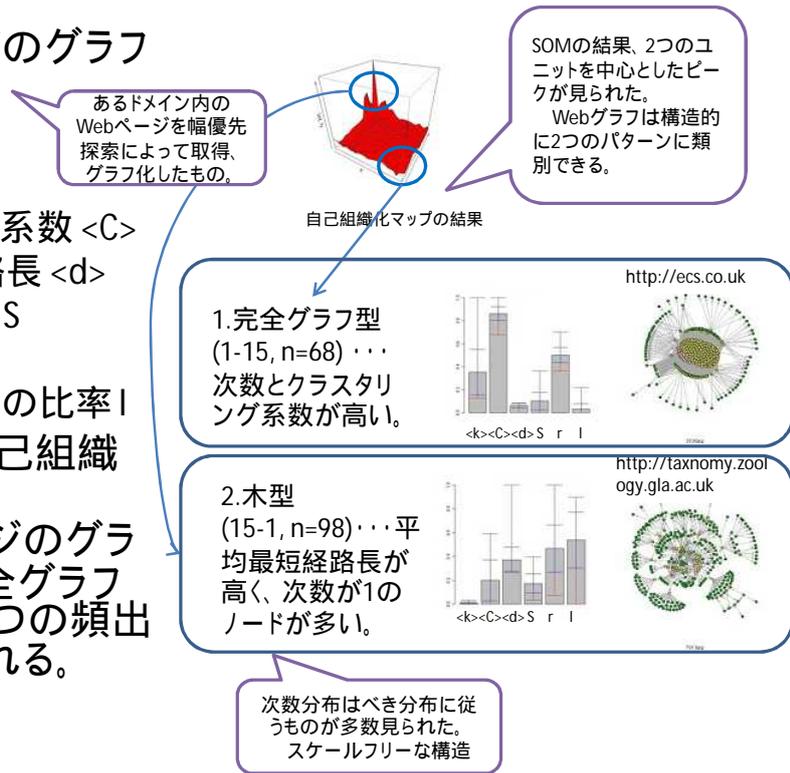
長所

1. グラフのマクロな特徴量を用いることにより、グラフの構造を理解できる。
2. 計算量が少ない。



実験と結果

- 対象: Webページのグラフ (N=1090)
- 特徴ベクトル:
 1. 平均次数 $\langle k \rangle$
 2. クラスタリング係数 $\langle C \rangle$
 3. 平均最短経路長 $\langle d \rangle$
 4. コミュニティ数 S
 5. 次数相関 r
 6. 次数1のノードの比率 l
- アルゴリズム: 自己組織化マップ(15 × 15)
- 結論: Webページのグラフにおいては完全グラフ型とツリー型の2つの頻出する構造が見られる。



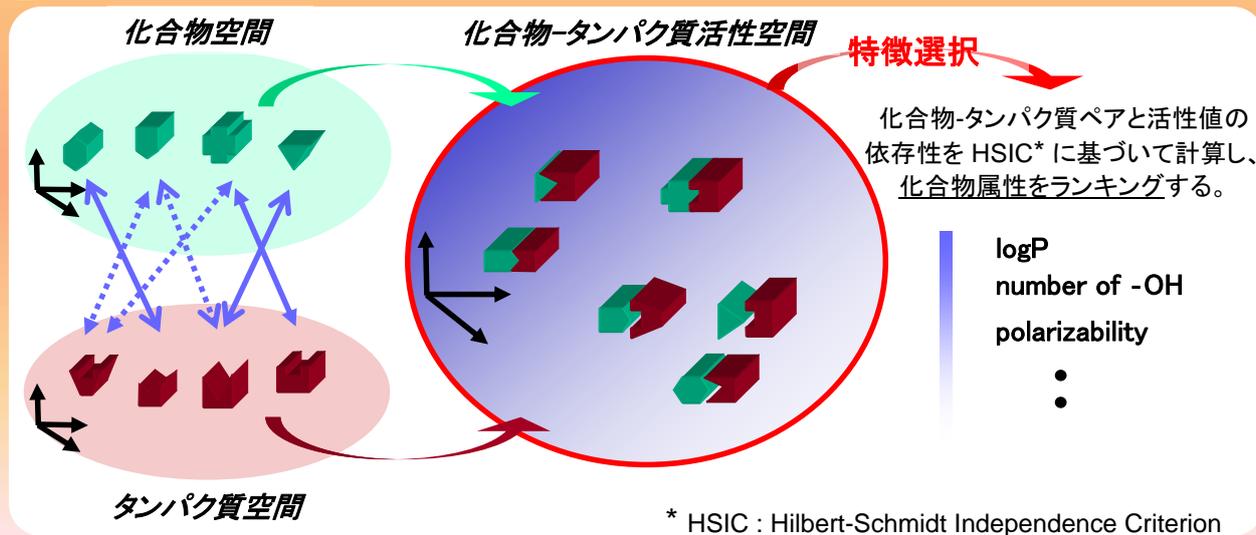
P037 化合物-タンパク質活性空間における特徴選択

新島 聡, 奥野 恭史
 京都大学大学院薬学研究科
 {niiijima,okuno}@pharm.kyoto-u.ac.jp

目的

カーネル関数を介して構成される化合物-タンパク質活性空間において特徴選択を可能とする効率的なアルゴリズムを提案し、活性予測に寄与する化合物属性を同定する。

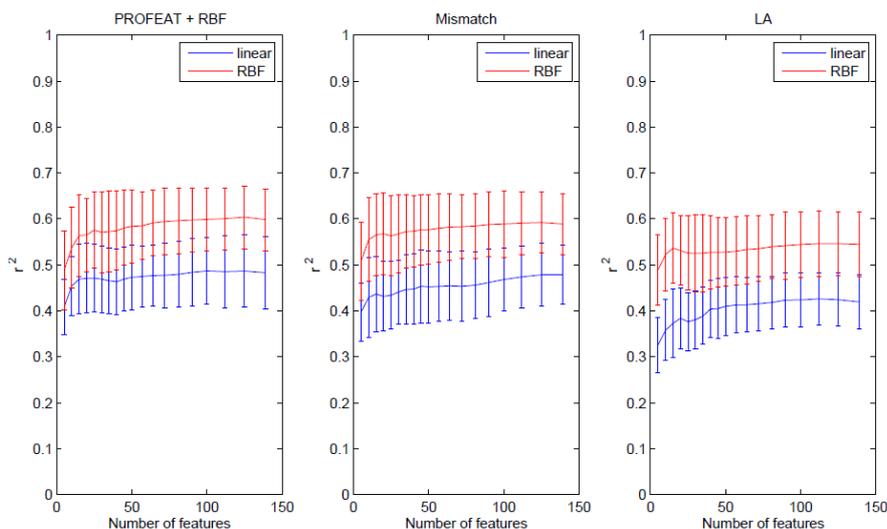
方法



実験結果 CYP阻害活性予測

- ◆ 活性データ(化合物-タンパク質ペア)数 = 798
- ◆ タンパク質(CYP)数 = 14 ◆ 化合物数 = 371 ◆ 化合物属性数 = 139

SVR の予測性能



化合物 : linear / RBF kernel
 タンパク質 : PROFEAT+RBF / Mismatch / LA (Local Alignment) kernel
 化合物-タンパク質ペア : tensor product kernel

化合物属性のランキング

rank	description
1	AlogP
2	MlogP ²
3	number of benzene-like rings
4	aromatic ratio
5	number of 6-membered rings
6	number of rotatable bonds
7	MlogP
8	mean absolute charge (charge polarization)
9	AlogP ²
10	number of ethers (aliphatic)

P043 大幾何マージン最小分類誤り学習法

渡辺 秀行¹ 片桐 滋² 山田 幸太² マクダーモット エリック³
 中村 篤³ 渡部 晋治³ 大崎 美穂²

1 情報通信研究機構 MASTARプロジェクト 音声コミュニケーショングループ
 2 同志社大学大学院 工学研究科 情報工学専攻
 3 日本電信電話株式会社 NTTコミュニケーション科学基礎研究所

E-mail: hideyuki.watanabe@nict.go.jp

最小分類誤り (MCE) 学習

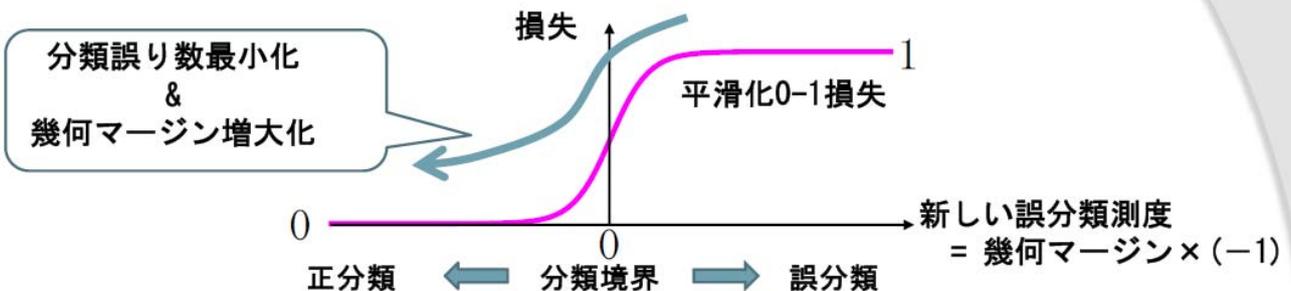
- 標本集合の分類誤り数の最小化を直接的に追求する.
- 多様な判別関数に適用可能.
- 有限の標本集合を用いた学習では, 究極の設計目標である, 分類誤り確率の最小化にならない.

+ 判別関数の一般形に対する幾何マージンの導出とその増大化を伴うMCE学習の提案.

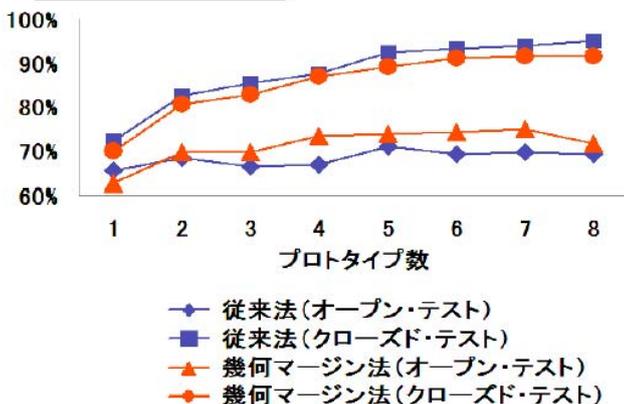
幾何マージン

- 分類決定境界と, 境界に最も近い学習標本との距離.
- 値を大きくとれば, 頑健性 (汎化能力) 向上が期待できる.

提案手法: 幾何マージン増大化を伴うMCE学習



実験結果



- 複数プロトタイプを用いた分類器による実装.
- 提案手法 (幾何マージン法) と, 従来の誤分類測度を用いたMCE学習法 (従来法) との比較.
- Leave-One-Out法による認識率の算出.



- 幾何マージン法の性能が従来法と比較して, クローズド・テストでは低いが, オープン・テストでは (プロトタイプ数が1個のときを除いて) 安定して高い.
- 幾何マージン法の未知標本に対する頑健性の高さを実証.

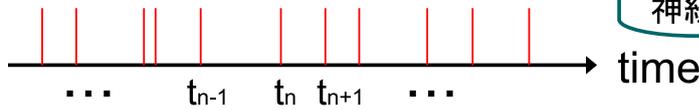
P046 点過程を特徴付ける統計量の時間変化を推定する

下川 文明, 篠本 滋

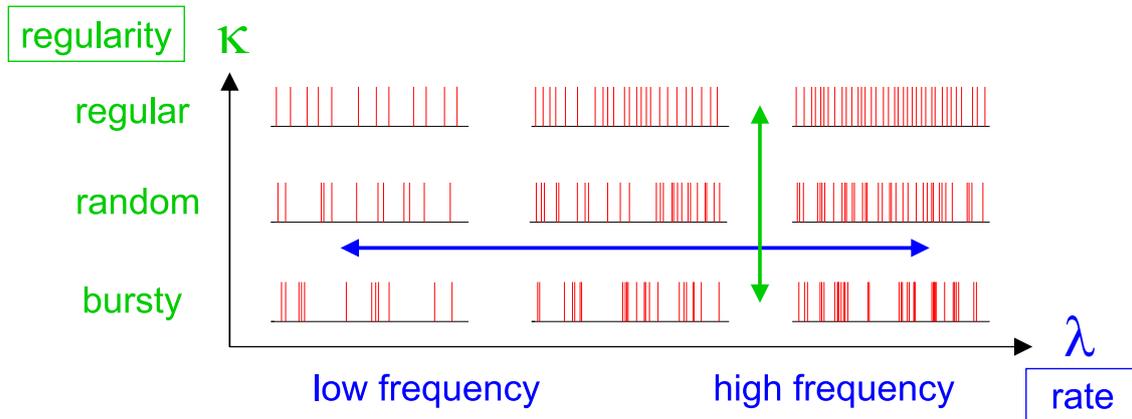
京都大学理学研究科 物理学・宇宙物理学専攻
shimokawa@ton.scphys.kyoto-u.ac.jp

- イベント発生時刻のデータ

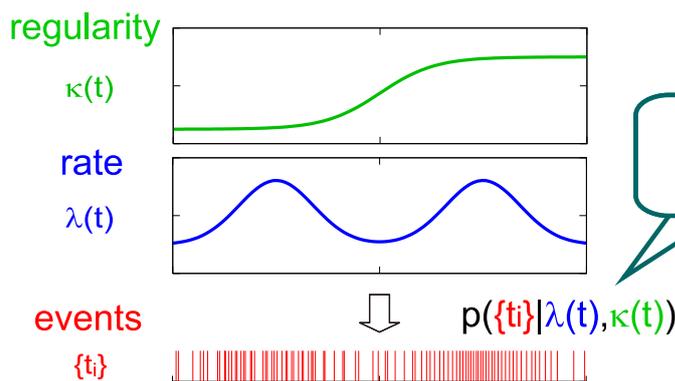
地震, webアクセス, 故障, 神経スパイク, 事件, etc.



- 発生率(rate)と不規則性(irregularity or regularity)によりうまく特徴付けできる (←情報幾何的に直交する)



- データから発生率と不規則性の時間変化を読み取ることのできるベイズ推定法を考案した



gamma process

Gaussian process prior

$$p(\lambda(t)) \propto \exp \left[-\frac{1}{2\gamma_\lambda^2} \int_0^T \left(\frac{d\lambda(t)}{dt} \right)^2 dt \right]$$

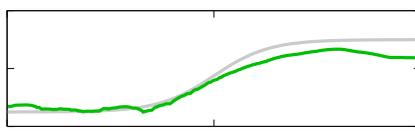
$$p(\kappa(t)) \propto \exp \left[-\frac{1}{2\gamma_\kappa^2} \int_0^T \left(\frac{d\kappa(t)}{dt} \right)^2 dt \right]$$

Bayesian inference

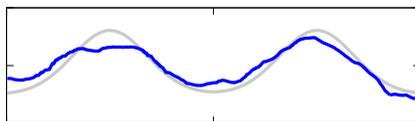


$$p(\lambda(t), \kappa(t) | \{t_i\}) = \frac{p(\{t_i\} | \lambda(t), \kappa(t)) p(\lambda(t)) p(\kappa(t))}{p(\{t_i\})}$$

$\hat{\kappa}(t)$



$\hat{\lambda}(t)$



hyperparameter selection

$$(\gamma_\lambda, \gamma_\kappa) = \operatorname{argmax}_{(\gamma_\lambda, \gamma_\kappa)} p(\{t_i\}; \gamma_\lambda, \gamma_\kappa)$$

T. Shimokawa and S. Shinomoto,
Neural Computation (2009) 21:1931-1951

P050 生物学情報への機械学習解析の応用 (Toxicogenomicsへの展開)



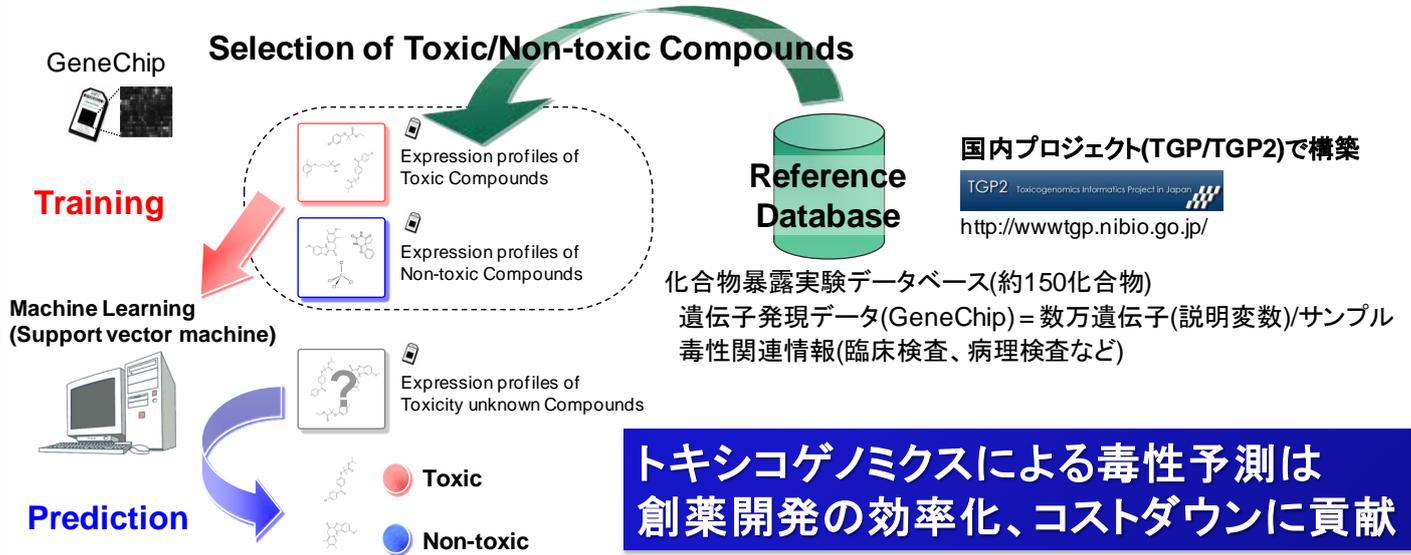
武藤裕紀*, 松下智哉, 芦原基起

*中外製薬株式会社 研究本部 創薬資源研究部

e-mail : mutohrn@chugai-pharm.co.jp

概要

機械学習をトキシコゲノミクスデータ解析へ応用し、データ前処理方法の選択による予測モデルへの影響検討および化合物の毒性予測を行った



結果

データ前処理方法として、①GeneChip数値化手法(Fig. 1)、②プローブセット選択手法の予測モデルへの影響を検討した

最適な手法の選択が予測精度の向上に重要であることが確認された

構築した予測モデルで、151化合物の毒性を予測した(Table 1)

機械学習を応用することで化合物の毒性を予測できる可能性が示唆された

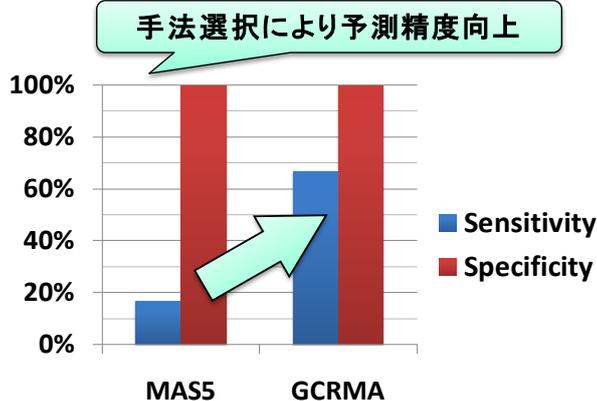


Fig. 1 GeneChip 数値化手法の影響検討

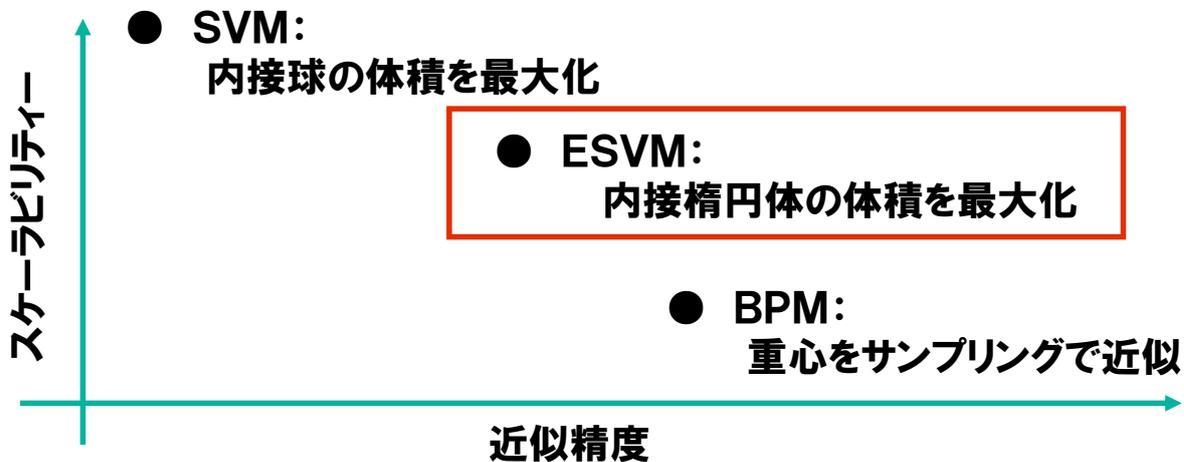
毒性有り化合物の多くを正しく分類

Table 1 毒性有りと予測された化合物

Compound // Time point & Dose	4 day			8 day			15 day			29 day		
	L	M	H	L	M	H	L	M	H	L	M	H
acetamidofluorene						T		T	T	T	T	T
allyl alcohol												T*
cisplatin												T
colchicine				T								
ethambutol						T				T		
lomustine										T		T*
lornoxicam				T			T					NA
meloxicam							T					NA
methapyrilene						T				T		T*
monocrotaline							T		T	T		T
naphthyl isothiocyanate							T			T		T*
naproxen				T						T		NA
nitrosodiethylamine							T		T	T		T
phalloidin	NA	NA		T	NA	NA	T	NA	NA	NA	NA	NA
thioacetamide				T						T		T*

L:Low dose, M:Middle dose, H:High dose, T:Toxic(predicted), NA:Not Available
Red colored:Proliferation, bile duct was observed
*Training sample

分類モデルの設計: バージョン空間の“重心”をどうやって見つけるか?



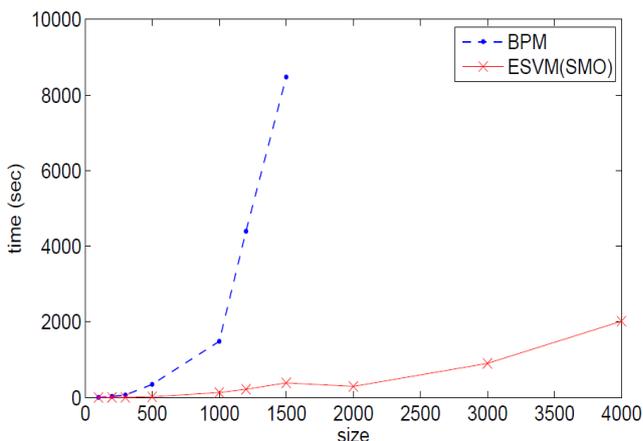
SVM: Support vector machine
 ESVM: Ellipsoidal support vector machine
 BPM: Bayes point machine

実験結果

汎化性能

DATA SET	SVM	BPM	E-SVM
THYROID	4.96 (.24)	4.24 (.22)	4.42 (.25)
HEART	25.86 (.40)	22.58 (.33)	20.87 (.32)
DIABETES	33.87 (.21)	31.06 (.22)	29.68 (.24)
WAVE	13.19 (.12)	12.02 (.08)	11.59 (.07)
BANANA	16.24 (.14)	13.70 (.10)	12.76 (.08)
WISC-BC	4.22 (.13)	2.56 (.10)	3.28 (.12)
BUPA	37.04 (.39)	34.5 (.38)	32.71 (.35)
GERMAN	30.07 (.22)	27.16 (.24)	26.34 (.27)
BREST	35.17 (.51)	33.04 (.48)	31.96 (.51)
SONAR	14.90 (.38)	16.26 (.36)	16.87 (.38)
IONO	7.94 (.25)	11.45 (.25)	5.92 (.21)

スケーラビリティ



ESVMの分類精度はBPMとほぼ同じ、計算効率は劇的に改善

P57 確率伝搬法による確率的画像処理 における統計的性能評価

東北大学大学院 情報科学研究科

片岡 駿, 安田 宗樹, 田中 和之

{xkataoka, muneki, kazu}@smqip.is.tohoku.ac.jp

・確率的画像処理の統計的性能評価を確率伝搬法を用いて行う。

・関連研究:

Y.Kabashima et alによる誤り訂正符号,

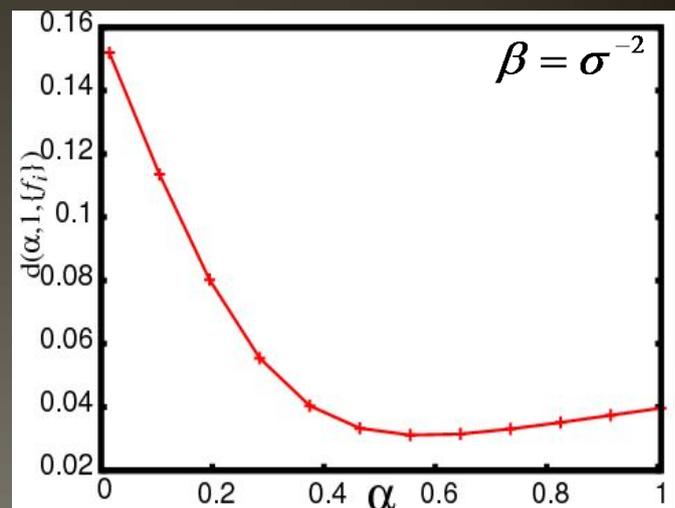
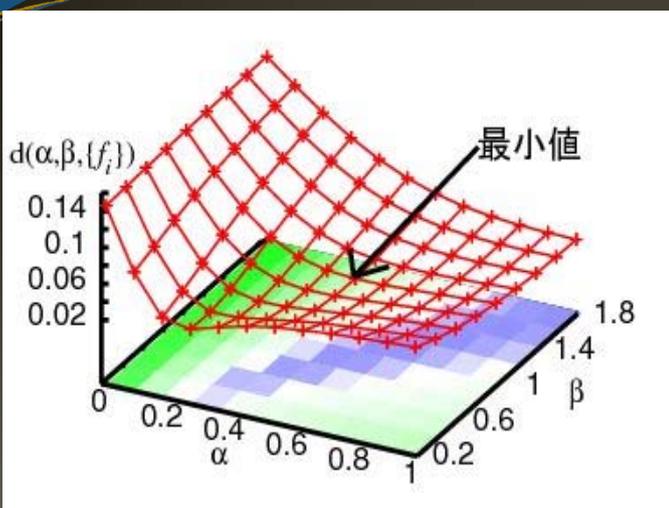
T.TanakaによるCDMA,

H.Nishimori et alによる画像処理

におけるレプリカ法を用いた統計的性能評価

・本講演では完全グラフ, 疎なランダムグラフではなく
正方格子上のグラフィカルモデルを扱う

実験結果



原画像を与えたときの確率的画像処理の統計平均

α, β : ハイパラメータ

$d(\alpha, \beta, \{f_i\})$: 原画像と修復画像の間の
ハミング距離の平均

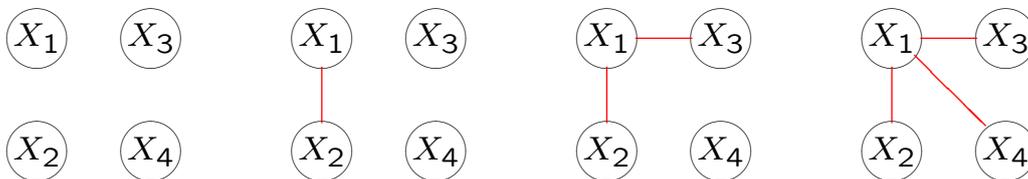


#64 Chow-Liu アルゴリズムの一般化と木の複雑さを考慮した修正版について

石田 悠、鈴木 讓 (大阪大学, 2009年10月20日)

Chow-Liu アルゴリズム: Markov ネットの分布を **K-L 情報量最小**の木で近似
 ループが生成されない限り、**相互情報量最大**の2頂点 (**確率変数**) を辺で結ぶ
 (サンプルから始まれば、**学習**にも使える)

$$I(X_1, X_2) > I(X_1, X_3) > I(X_2, X_3) > I(X_1, X_4) > I(X_2, X_4) > I(X_3, X_4)$$



- 尤度最大の木 $\iff I(X_i, X_j)$ 最大
- **MDL 最小**の木 $\iff I(X_i, X_j) - \frac{1}{2}(\alpha_i - 1)(\alpha_j - 1) \log n$ 最大
 (α_i : X_i のとる値の数)、木の複雑さも考慮 (Suzuki, 1993)

問題意識: **学習** というと、**有限** とか、**連続** とか、特殊なケースばかり

- 一般の確率変数に対する、Chow-Liu アルゴリズム
- (有限でも連続でもない) 一般的な相互情報量の定義
- Suzuki の修正版を、正規分布の確率変数を持つ場合に適用
- Suzuki の修正版を、有限と正規分布の確率変数が混在する場合に適用

課題: Suzuki の修正版は、究極的にどこまで一般化されるのか。

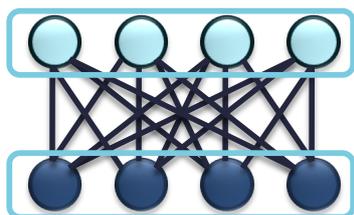
70: KULLBACK-LEIBLER IMPORTANCE ESTIMATION PROCEDURE を用いた RESTRICTED BOLTZMANN MACHINE の学習アルゴリズム

桜井 哲治, 安田 宗樹, 田中 和之

東北大学大学院情報科学研究科応用情報科学専攻

KEYWORDS : deep belief network, restricted Boltzmann machine, variational bound, Kullback-Leibler Importance Estimation Procedure

Restricted Boltzmann Machine (RBM)



隠れ素子 • 可視素子と隠れ素子の層からなる二部グラフの構造をもつ確率モデル

可視素子 • Deep belief network の構成要素

Roux and Bengio (RB) Algorithm

Deep belief network に対して有効なRBMの学習アルゴリズム



計算量の問題から厳密な計算は困難！！

提案アルゴリズム

Kullback-Leibler Importance Estimation Procedure (KLIEP) をもちいた RBアルゴリズムの近似アルゴリズム

RBアルゴリズム

高コスト

KLIEP

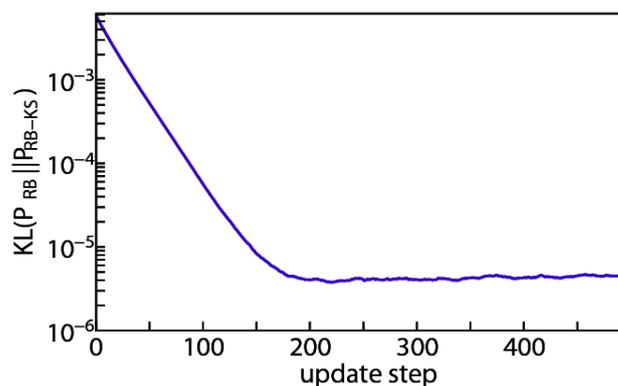
提案アルゴリズム

低コスト！！

実験結果

縦軸: RBアルゴリズムの結果と提案アルゴリズムの結果との間のKL情報量

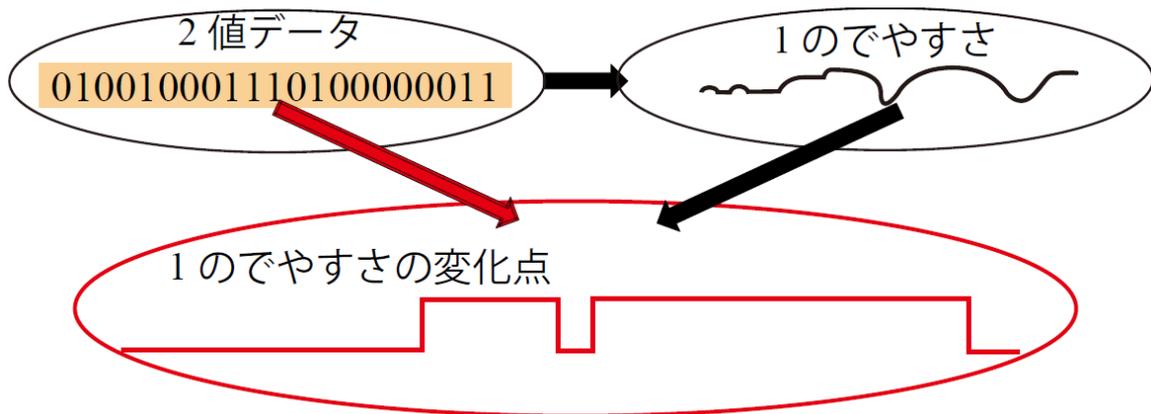
横軸: 提案アルゴリズムのアップデートステップ



P072 非定常2値時系列データから隠れた構造を読み取る
 — 神経科学データへの応用 —
 瀧山健 (東大新領域), 岡田真人 (東大新領域, 理研BSI)

用いた解析手法: スイッチング状態空間モデル(Gharamani & Hinton 2000)
 変分ベイズ法(Attias 1999)

2値時系列データから変化点を検出できるか？



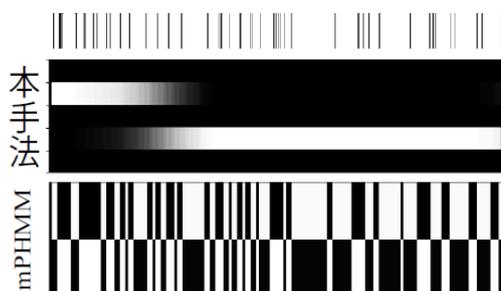
2値データの例...地震, 保険, 経済時系列, 生存解析, **神経科学**

脳は10の10乗個のニューロンから構成. 観測は局所的. 変化点⇔脳の状態？



限られたニューロンの活動データから,
 我々はどれだけ脳全体の状態を読み取ることができるか？

実データ解析結果



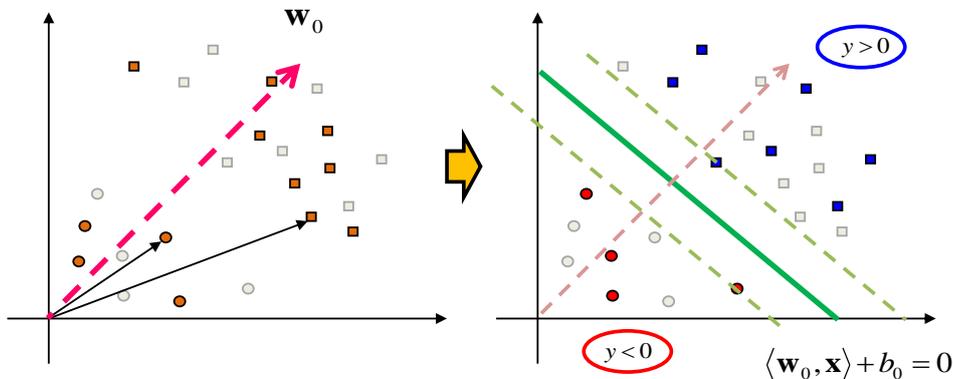
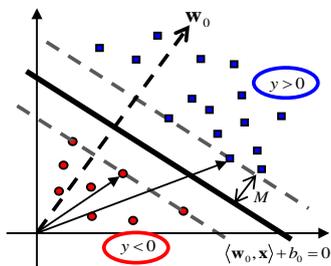
P075 部分空間SVMのための交差部分空間学習

井之上 直矢
Naoya INOUE

山下 幸彦
Yukihiko YAMASHITA



部分空間SVM



従来のSVM

識別超平面の

★ 法線ベクトル ★ 閾値



同じ学習データから

SVMを部分空間法で拡張する



部分空間SVM

識別超平面の法線ベクトルを部分空間で制限する

参照: K-fold 交差検定

サンプルを K 分割

$K-1$ のセットを用いてモデルを構築 → 残った1つで評価

以下、繰り返し

例: データセットを2つに分けた場合



識別

誤認識率を計算

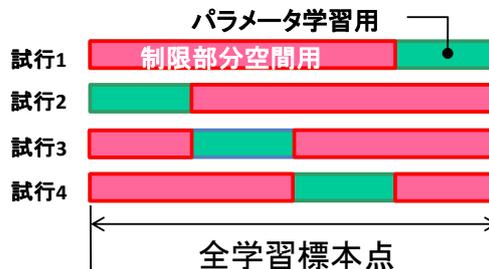
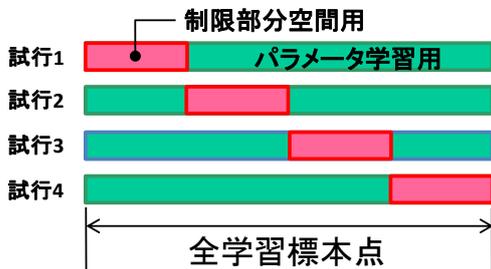
これを部分空間SVMに組み込む



交差部分空間学習

交差部分空間学習

例: $K=4$



実験・結果

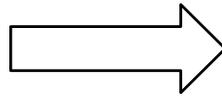
★ 2クラス識別用の13データセット

★ 様々な分割例

本手法の有用性を示した

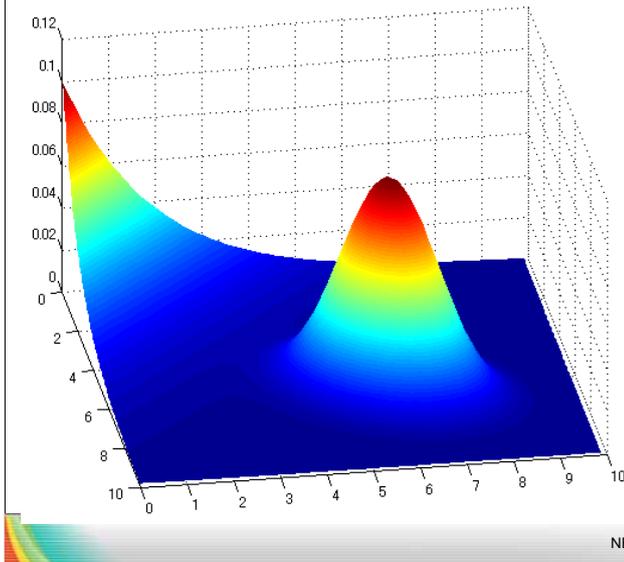
異種混合モデル選択

- コンポーネントの数
- コンポーネントの種類

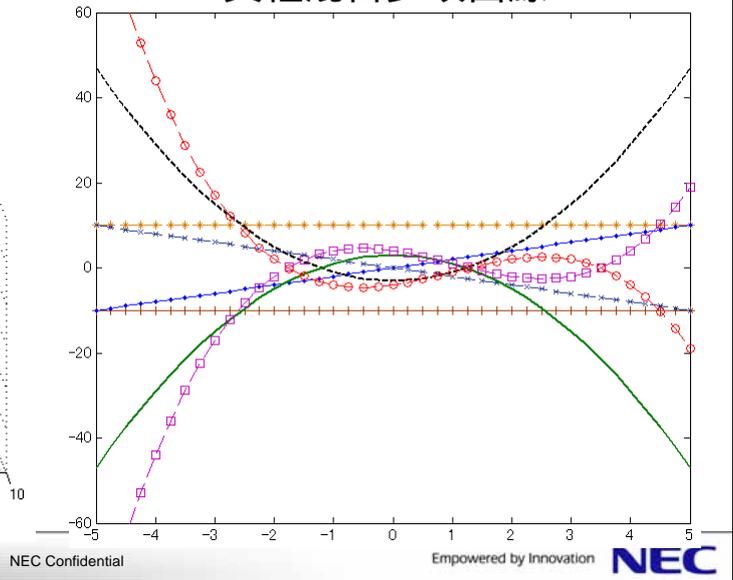


モデルの候補数が膨大
 (コンポーネントの組合せ)

異種混合分布 (正規 + 指数)



異種混合多項曲線



期待情報量基準最小化法 (EMIC)

- 完全変数の期待情報量基準を最小化

	EMIC	EM
目的関数	期待情報量基準	期待対数尤度
各EMでの最適化	コンポーネント種類 コンポーネントパラメータ	コンポーネントパラメータ
正当性	情報量基準の単調減少	対数尤度の単調増加

利点

- 任意のコンポーネント候補を扱う事が可能
- コンポーネント候補数、コンポーネント数に対して線形時間で最適化

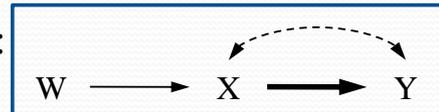
Extending the Use of Instrumental Variables for the Identification of Direct Causal Effects in SEMs

- Problem: Identify direct causal effect $X \rightarrow Y$ in linear Structural Equation Models (SEMs)

- Previous result: Use **instrumental variable** W , if:

- Graphical criteria is satisfied, e.g.:

- W is **not a descendant of X or Y**



- Direct effect $X \rightarrow Y = \sigma_{wy.z} / \sigma_{wx.z}$
- However, we cannot identify some direct effects $X \rightarrow Y$ using this graphical criteria
- Question: Can we use W as an instrumental variable, even if W is **a descendant of X or Y** ?

Hei Chan (AIST Japan), Manabu Kuroki (Osaka University)

Extending the Use of Instrumental Variables for the Identification of Direct Causal Effects in SEMs

- New result: W , **a descendant of X** , can be used as an “**path-specific instrumental variable**”, if:

- New graphical criteria is satisfied, e.g.:

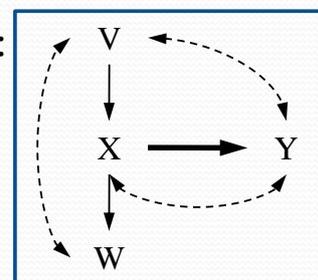
- Total effect $X \rightarrow W$ is identifiable: τ_{wx}

- Direct effect $X \rightarrow Y =$

$$(\sigma_{wy.z} - \sigma_{yx.z} \tau_{wx}) / (\sigma_{wx.z} - \sigma_{xx.z} \tau_{wx})$$

- As a result, we can identify some direct effects $X \rightarrow Y$ using this new graphical criteria

- Future work: Solution of W , **a descendant of Y** , as an “path-specific instrumental variable”



Hei Chan (AIST Japan), Manabu Kuroki (Osaka University)

P080: 独立性の一般化に基づく統計モデルの拡張

藤本 悠 (青山学院大学), 村田 昇 (早稲田大学)
 yu.fujimoto@it.aoyama.ac.jp

【背景】

パラメトリックな統計モデル

- 独立性
- 線形性
- 正規性
- ...

独立: $p(X, Y) = p(X) \times p(Y)$

周辺分布の情報のみで同時分布が再現できる

非独立: $p(X, Y) \neq p(X) \times p(Y)$

周辺分布の情報のみでは同時分布が再現できない

【通常の独立性】

周辺分布の積

$$p(X) \times p(Y) = \exp(\log(p(X)) + \log(p(Y)))$$

が同時分布 $p(X, Y)$ に対応するかどうか



【一般化した独立性】

周辺分布の (一種の) 積

$$p(X) \times p(Y) = u(\xi(p(X)) + \xi(p(Y)) - c)$$

が同時分布 $p(X, Y)$ に対応するかどうか



●Kullback-Leibler情報量

$$D_{KL}(p, q) = \sum (p(x)(\log(p(x)) - \log(q(x))))$$

一般的な統計的推論で用いられる推定指標



●Bregman情報量

$$D_U(p, q) = \sum (U(q(x)) - U(p(x)) - p(x)(\xi(q(x)) - \xi(p(x))))$$

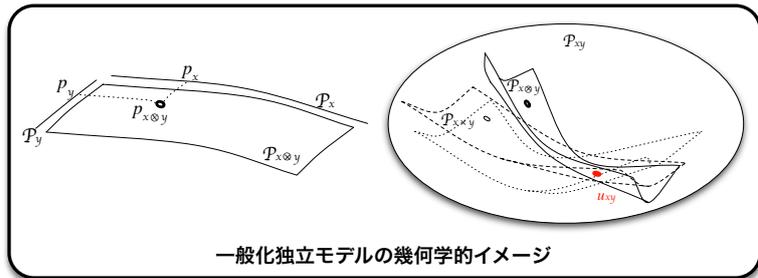
任意の関数 $U(\cdot)$, $\xi(\cdot)$ 等を用いることでロバスト性などの特殊な性質の実現が可能な推定指標

IBIS2009 ポスタープレビュー

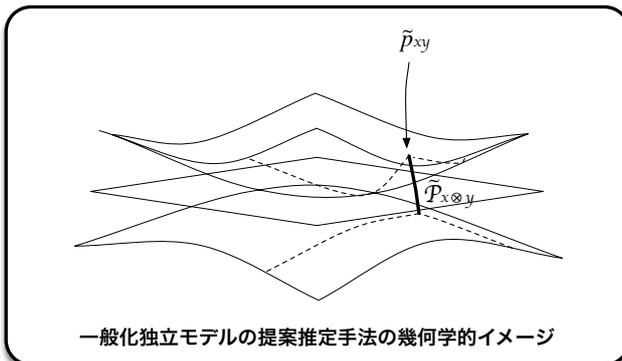
【モデルの拡張】

一般化独立性の考え方に基づいて、弱い特殊な依存関係をシンプルに記述した統計モデルを構築する

- 独立モデル
- NB (Naive Bayes)
- BN (Bayesian Network)
- 対数線形モデル
- ...



【モデルの推定法に関する議論】



【その他の話題】

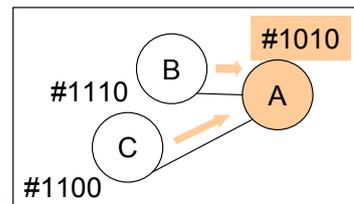
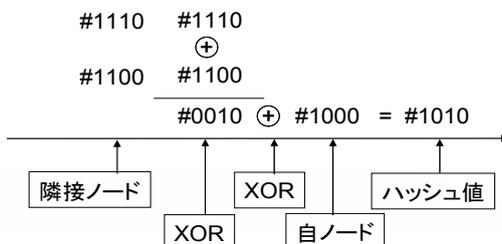
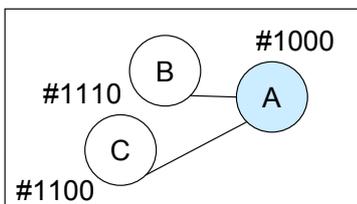
- 他の依存関係表現との比較
- Copulaとの関係
- 数値実験結果
- ...

IBIS2009 ポスタープレビュー

2

目的: 数千ノードグラフ用の高速なカーネルの構築

近傍ハッシュカーネル



(1) ビットラベル化

- ・ 離散ノードラベルを固定長ビット列で表現

(2) 近傍ハッシュラベル計算

- ・ XORなど論理演算で効率的に隣接ノード情報を集約する

(3) ラベル更新

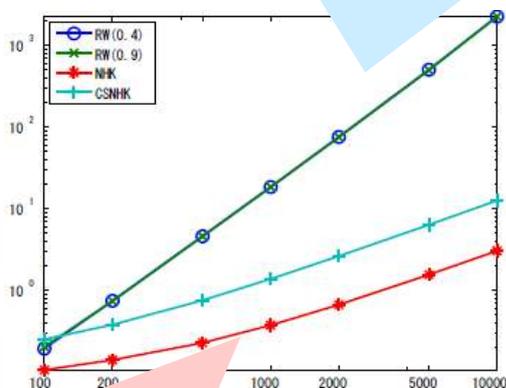
- ・ 周辺ラベルの分布を反映したラベルとなる

(4) 全ノードに適用し、更新ラベルの一致割合をカーネル値とする



人工グラフに対する計算量

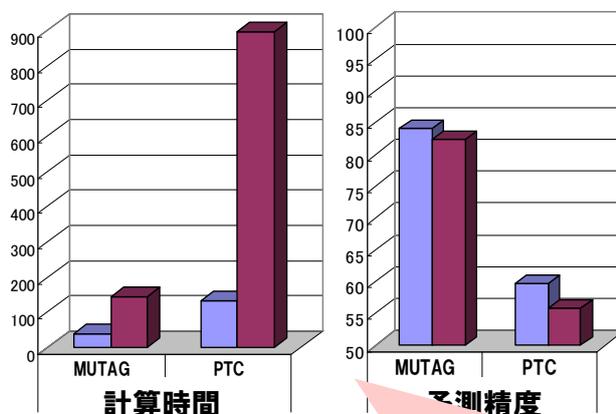
既存カーネル: グラフサイズに二乗



近傍ハッシュカーネル: サイズに線形

ベンチマークの性能と計算時間

■ 近傍ハッシュカーネル ■ ランダムウォークカーネル



計算時間: 小 & 精度: 良

近傍ハッシュカーネル: サイズに線形の計算量 & 高性能



#86 Hannan-Quinn の命題は、線形回帰でも、
ガウス型 Bayesian ネットワークの構造推定でも正しい
鈴木 讓 (大阪大学, 2009年10月20日)

n 個の例から、**情報量基準** = 経験的エントロピー + パラメータ数 $\times d_n$
($d_n = 1 \implies$ AIC, $d_n = \log n/2 \implies$ MDL/BIC)
最小のモデルを選択したい
 $n \rightarrow \infty$ で正しいモデルを推定したい (**一貫性**)

d_n が小さすぎると、一貫性は満足されない

- 自己回帰移動平均 (ARMA) の次数推定 $d_n \geq \log \log n$ (Hannan-Quinn, 1979)
- 条件付確率の状態分割 (有限型 BN の構造推定) $d_n \geq \log \log n$ (Suzuki, 2006)

1

- **重回帰分析** の変数選択 (ガウス型 BN の構造推定) ?

$$Y = \sum_{j=1}^m \alpha_j X_j + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2)$$

n 個の例 $\{(y_i, x_{i,1}, \dots, x_{i,m})\}_{i=1}^n$ から変数 $\{X_j | \alpha_j \neq 0\}$ を推定

重回帰分析でも $d_n \geq \log \log n$ を証明することができた (**重複対数の法則**)

従来は、 $\frac{d_n}{\log \log n} \rightarrow \infty$ (十分条件にすぎない)

モデル選択の誤り率: 条件付確率の状態分割 (Suzuki, 2006) と導出が似ている

過学習 多すぎた変数の数を自由度とする χ^2 分布にしたがう

未学習 (過学習以外の誤り) 0 に概収束

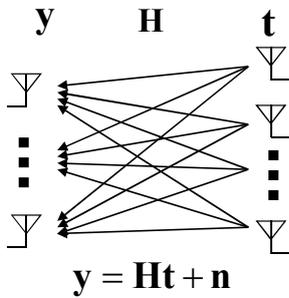
課題: **指数分布族** で共通しているのか? **有限 Markov の次数推定** でも同じ結果か?

P090 非線形プリコーディングの統計力学的解析

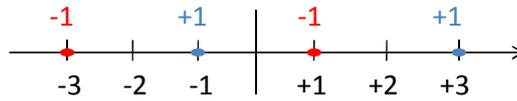
林 愛空, 樺島祥介
東京工業大学総合理工学研究科

モデル

MIMO通信



Tomlinson-Harashima precoding
(Tomlinson. '71, Harashima & Miyakawa. '72)



$$X(\mathbf{s}) = B_{s_1} \times B_{s_2} \times \dots \times B_{s_K}$$

$$B_{\pm 1} = \{\mp 1, \pm 3\}$$

$$\mathbf{t} = \arg \min_{\mathbf{x} \in X(\mathbf{s})} \frac{\mathbf{x}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{x}}{\text{送信電力}}$$

送信電力の低減

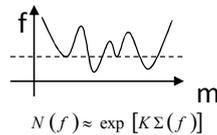
元信号

$$\mathbf{s} = \{\pm 1\}^K$$

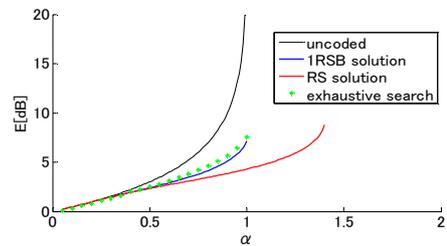
(同等の情報を持った
 2^K 個の候補から
最も送信電力の
小さい信号を送る)

統計力学的手法による解析

- 送信電力低減限界の評価
1RSB仮定でのレプリカ解析.
(解空間の多峰性を定量的に評価)



- TAP平均場法による符号化アルゴリズムの提案



Latent Dirichlet Allocationの量子アニーリング変分ベイズ学習 P095

○佐藤一誠¹, 栗原賢一², 田中宗¹, 宮下精二¹, 中川裕志¹
¹東京大学, ²Google

• Goal:

複数の状態を同時に取る

隠れ(離散)状態を量子状態として学習する



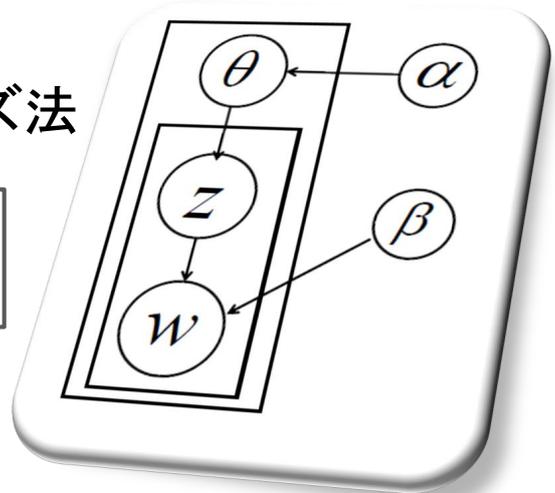
• Methodology:

量子アニーリング+ 変分ベイズ法

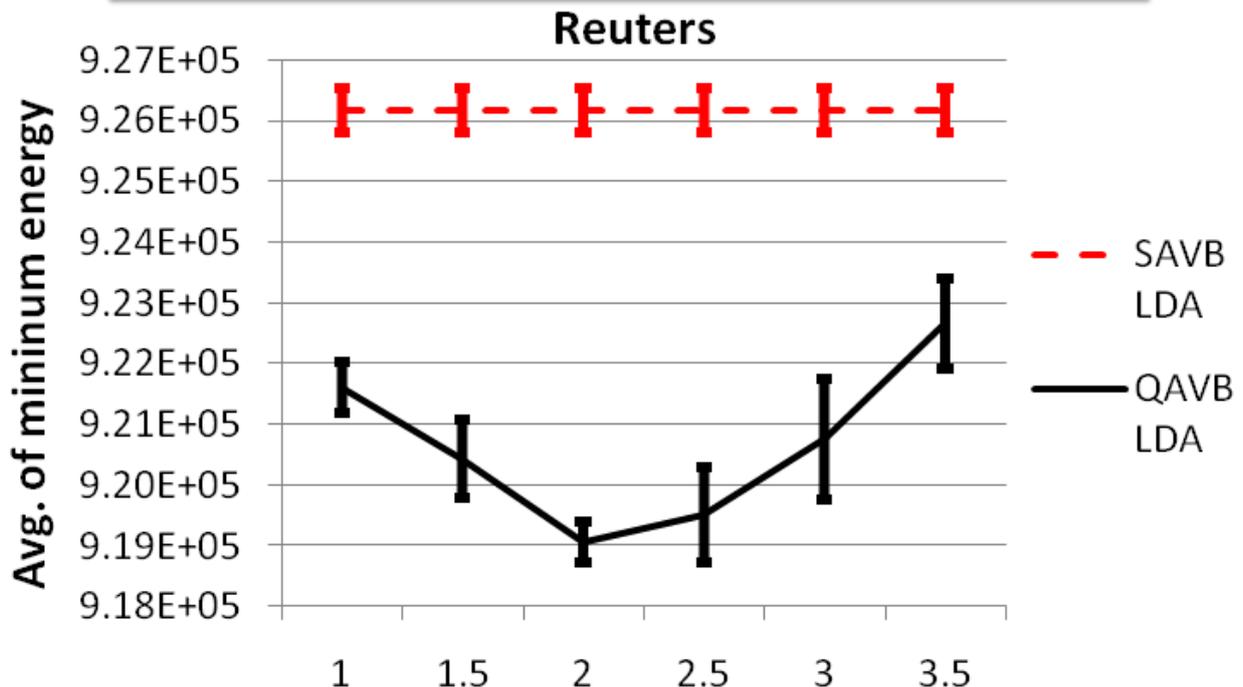
- 密度行列による隠れ状態表現
- 量子系から古典系への近似



Point



Performance of QAVB-LDA and SAVB-LDA



Γ_0

Quantum effect parameter



個性を考慮した周期的全身運動のオンライン予測

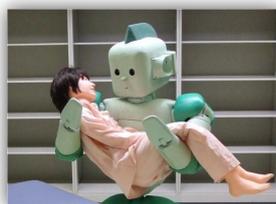
松原崇充^{1,2} 玄相昊² 森本淳²

1. 奈良先端科学技術大学院大学 2. ATR脳情報研究所

- ▶ 人全身周期運動のオンライン予測法を提案
 - ▣ 状態と個性を低次元変数で表現するコンパクトな生成モデル
 - ▣ 状態と個性を同時逐次推定するオンラインEMアルゴリズム

▶ 用途:

- ▣ ヒューマンロボットインタラクション
- ▣ ロボットスーツ
- ▣ 見真似学習

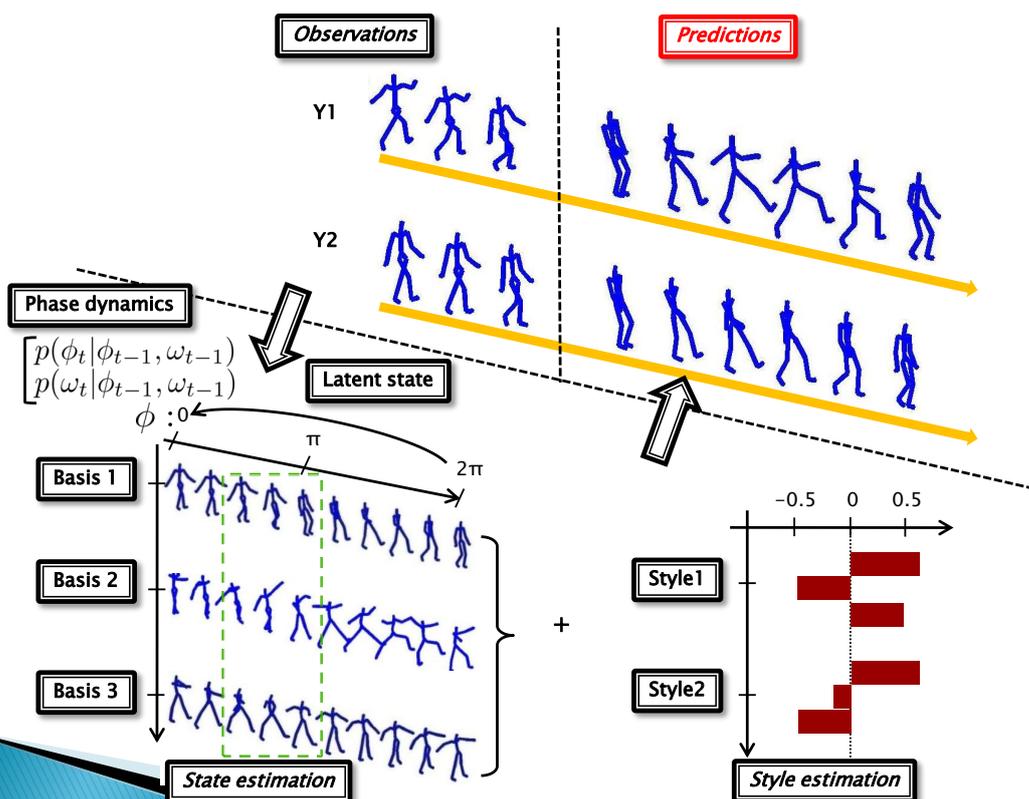


(Onishi et al.2007)



(Kawamoto et al.2003)

提案法の概要



第12回情報論的学習理論ワークショップ (IBIS 2009)

2009年10月19日(月)-21日(水) 九州大学医学部 百年講堂

ポスターセッションB

ディスカッショントラック

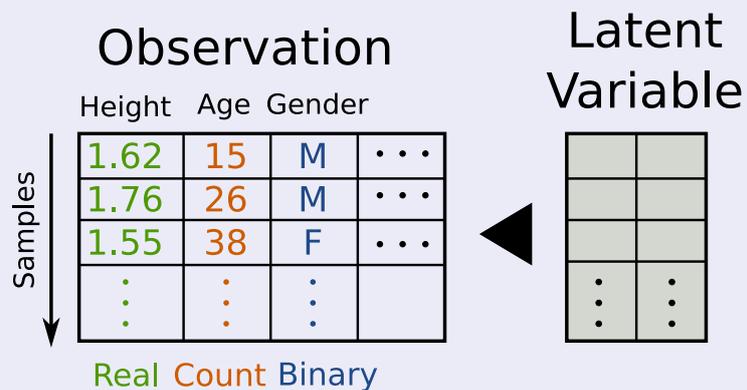


P012: Sparse Exponential Family PCA with Heterogeneous Attributes

K. Hayashi, T. Takenouchi, T. Shibata, K. Ikeda

Purpose

Low-dim. feature extraction from **mixed data types**

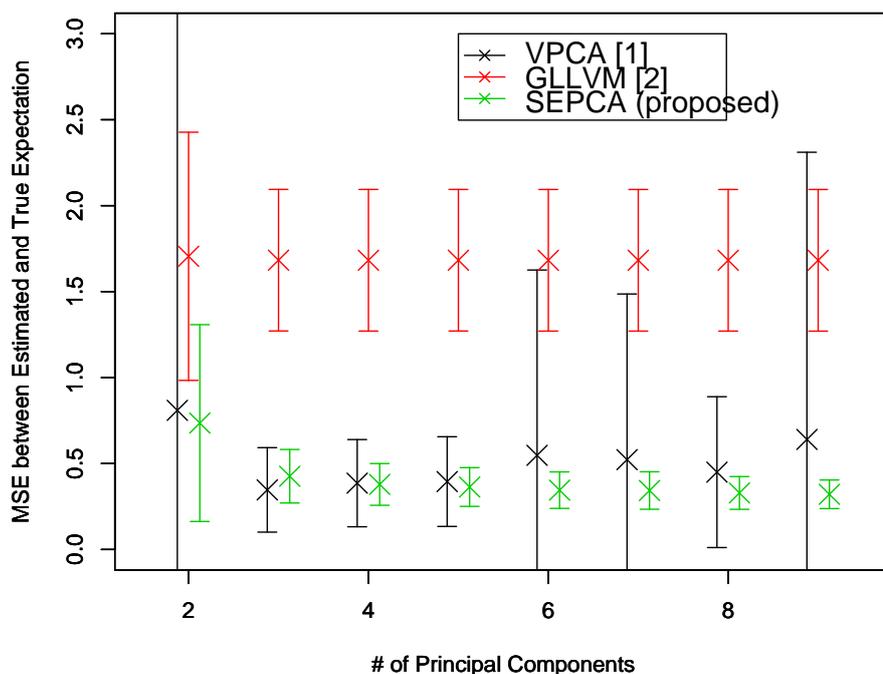


Approach

Modeling by **exponential family** with **sparse priors**

1/2

Synthetic Experiment Result



¹C. M. Bishop. Variational principal components. *ICANN '99*

²P. Huber *et al.* Estimation of generalized linear latent variable models. *J. Royal Stat. Soc. B*, '04

2/2

P15 複数タスクに対するタスク選択実施能力を有するニューラルネットワーク

八尾 泰洋

東京大学大学院工学系研究科

橋本 康弘

東京大学大学院工学系研究科

陳 Yu

東京大学大学院工学系研究科

大橋 弘忠

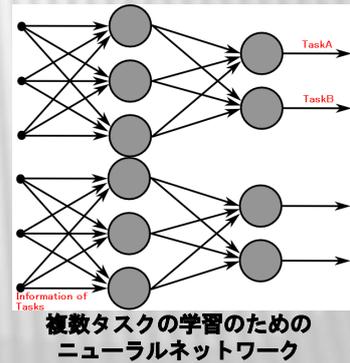
東京大学大学院工学系研究科

背景

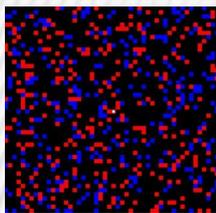
- 現実のタスクを機械に行わせる際には細分化、単純化が必要になることが多い。それに対し、生物はそのような複数のタスクを複雑なまま行うことができ、さらに以前の他のタスクでの経験を新しいタスクの学習に生かすことができる。
- ニューラルネットワークの複数タスク学習ではあるタスクが別のタスクの学習に生かされる例が報告されており、複数タスク学習によって上記のような生物の強みをニューラルネットワークに取り入れることができる。

目的、方法

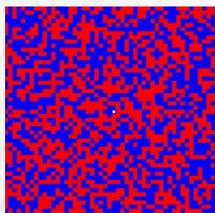
- 複数タスク学習の先行研究においてはニューラルネットワークの出力のから人為的にタスクに適合した出力が選ばれていた。
- 本研究では入力から判断してニューラルネットワークがどのタスクを行うかの選択を実行するようなモデルを考案した。
- 出力層をタスクごとに分割し人為的に出力を選択する方法と、タスクの種類という直接的な情報を入力として与えニューラルネットワークにタスクに応じた出力をさせる方法の2つの方法についてその成績を比較した。



実験結果



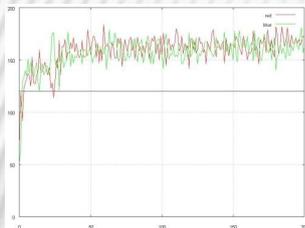
(a)



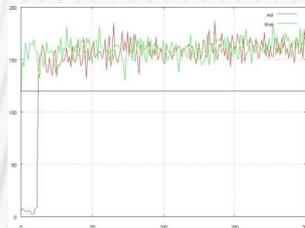
(b)

フィールド上で赤いトークンを集めるタスクと青いトークンを集めるタスクの2種類のタスクを同一のニューラルネットワークに課した。ニューラルネットワークがコントロールする駒（白点）は1ステップごとに上下左右のいずれかの方向に1マスずつ動いてトークンを集める。

(a),(b)ゲームを行うフィールド。赤点、青点がトークンを表す。(a)では各々300個ずつ、(b)では1250個ずつのトークンが配置してある。一代ごとにトークンの配置をランダムに生成しながらニューラルネットワークを進化させた。



(c)

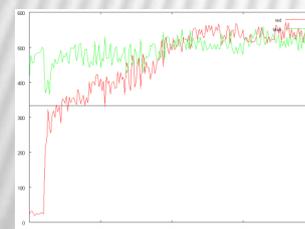


(d)

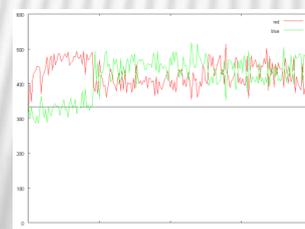
(c),(d),(e),(f)ニューラルネットワークの進化の様子。横軸が遺伝的アルゴリズムにおける世代、縦軸が集めたトークンの個数を示す。

(c),(d)初期状態で各種300個ずつのトークンを配置するフィールドで進化させた結果。(c)は人為的に出力を選択した場合の結果であり、(d)はニューラルネットワークにタスクに応じた出力をさせた場合の結果である。

(e),(f)初期状態で各種1250個ずつのトークンを配置するフィールドで進化させた結果。(e)は人為的に出力を選択した場合の結果であり、(f)はニューラルネットワークにタスクに応じた出力をさせた場合の結果である。



(e)



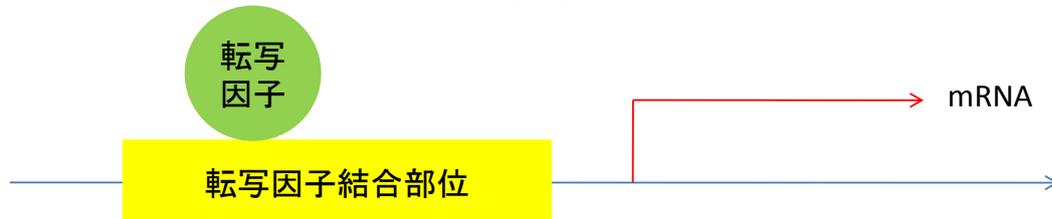
(f)

ニューラルネットワークは出力をタスクに適應させることができているが、トークンの個数が多い場合は人為的に出力を選択するほうが良い性能を示した。

How much is sequence predictive for gene expression?

白石 友一、長嶋剛史、岡田眞里子
理化学研究所 細胞システムモデル化研究チーム

- 遺伝子の転写制御メカニズムの解明への機械学習的なアプローチ
- 遺伝子発現に関係すると言われている因子
 - 転写因子結合配列
 - ヒストンの位置
 - エピジェネティックな情報(メチル化、アセチル化)



遺伝子の発現を配列から予測！

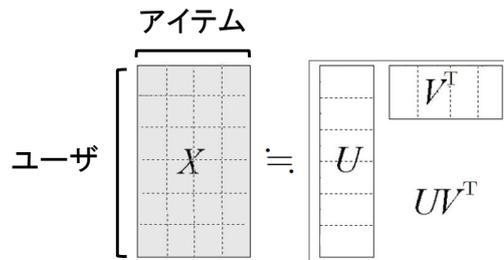
- 遺伝子上流配列を入力とし、遺伝子が発現が変動するか否かを出力とする判別機を構成
 - 上流配列のみで、どの程度発現の変動を予測できるか？の精度を判別率により検証。
 - 重要となる配列モチーフを選択し、生物学的な観点からの議論を加える。

P027 各要素が混合ガウス分布に従う行列に対する 行列因子化による欠損値予測

中村政義, 竹之内高志, 池田和司

奈良先端科学技術大学院大学

ユーザにグループが存在 } データに適した欠損値予測をしたい
グループで評価の傾向が異なる



ユーザ毎に属するグループを表す潜在変数を導入し
行列因子化法を拡張

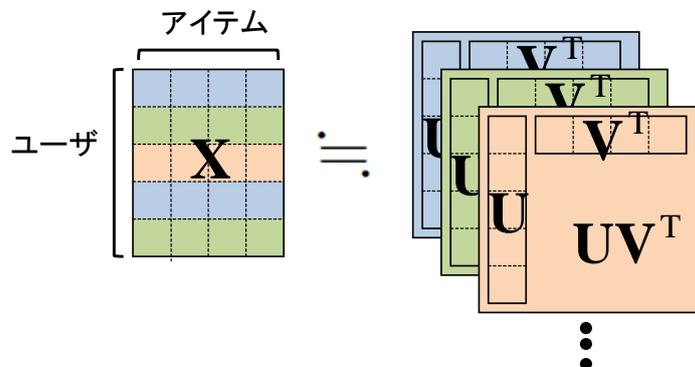
モデル化

$$p(x_{ij}) = \sum_{m=1}^M \pi_{mi} \mathcal{N}(x_{ij} | \mathbf{u}_{mi}^T \mathbf{v}_{mj}, \sigma_m^2)$$

x_{ij} : ユーザ i の商品 j への評価

m : グループ

π_{mi} : 混合係数



数値実験

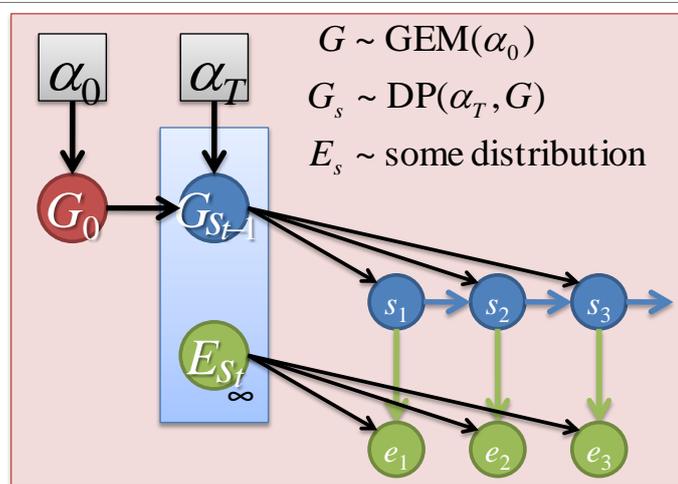
人工データ } に対して予測精度を比較
Netflix社の映画レーティングデータ

34. 階層状態無限隠れマルコフモデル

牧野 貴樹

東京大学 総括プロジェクト機構

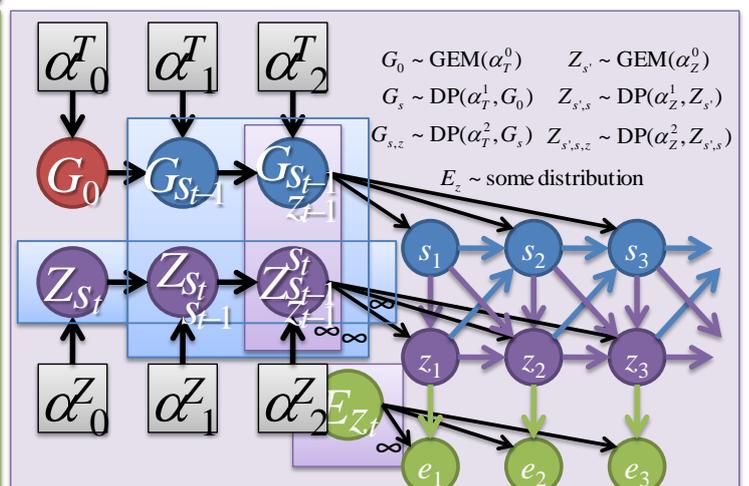
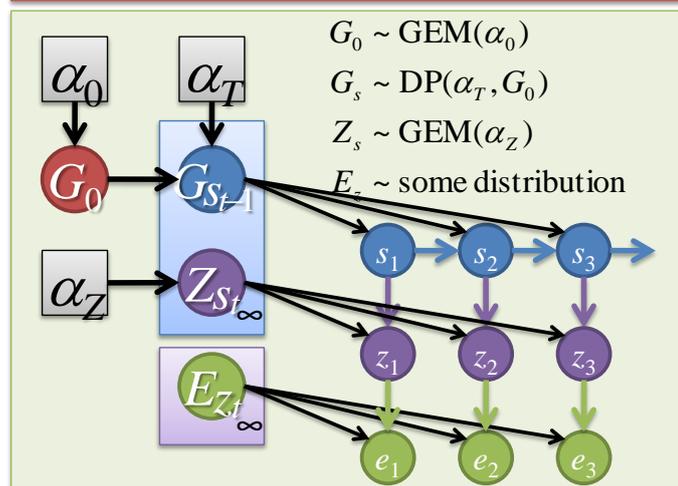
- 隠れマルコフモデル (HMM)とその状態のクラスタリングを、ノンパラメトリックベイズモデルで記述
 - HMMの学習は局所解に陥りやすい
 - ← いったん獲得した状態数を増減させることが難しい
 - 状態同士が確率的に類似するようなモデルであれば、状態数の増減に対する探索がしやすくなるのでは?
 - = 各状態が無数の子状態を階層的に含むモデル
- 階層 Chinese Restaurant Process を使った Particle Filter で実験 → 高速化、省メモリ化が確認できた



← 普通の HMM ベイズモデル

∠ HMM with Dirichlet Mixture

↓ 2階層状態 HMM ベイズモデル
(今回提案するモデルの1つ)



P042 階層パターンを持つ自己相関型連想記憶モデルのPCAによる解析

塩塚 丁二郎
早稲田大学

永田 賢二
東京大学

岡田 真人
東京大学

井上 真郷
早稲田大学

概要

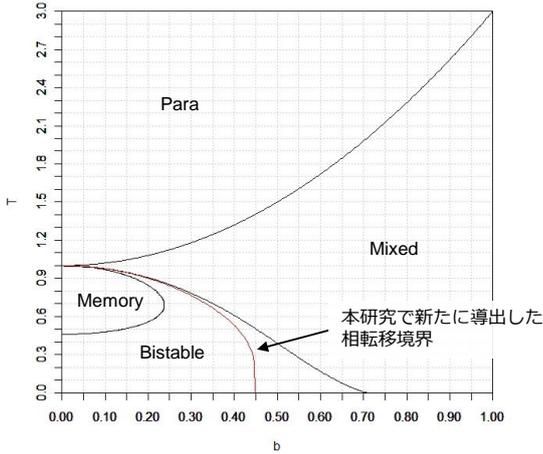
PCAを用いた高次元情報の可視化

解析対象

連想記憶モデル

あらかじめ埋め込んだ記憶を想起する脳の数理モデル。

→統計力学的手法によって詳細に解析されている。



パターン間の相関 b に関する相図
(親パターン: 1 子パターン: 3)

解析手法

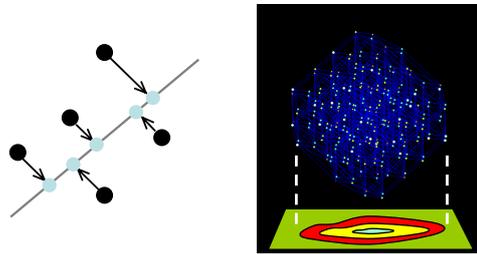
系全体のハミルトニアン

$$H = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} x_i x_j$$

ボルツマン分布

$$P(\mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{H(\mathbf{x})}{T}\right)$$

交換モンテカルロ法によってサンプリング



PCAを用いた可視化の様子
サンプリング点をヒストグラムで表す。

本研究で導出した相転移境界を挟む領域上でシミュレーションを行なった。
系の状態の変化（相転移）がどのように可視化されるのか。

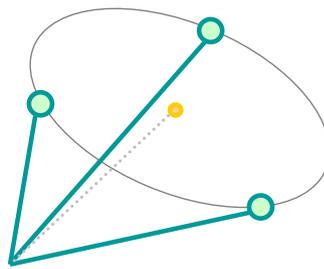
概念図

T_c 相転移点

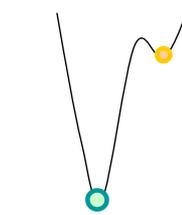
● 記憶パターン（子）

● 混合パターン（親）

$b < T_c$



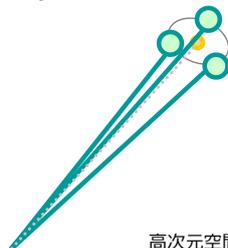
高次元空間上の様子



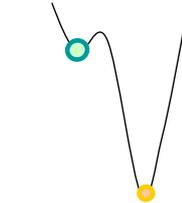
熱力学的エネルギーに関するイメージ図

記憶パターン優勢の領域では、記憶パターンが想起される。子パターン3個を記憶パターンとしたので、その反転パターンも含め6個のパターンの出現確率が高い。

$b > T_c$

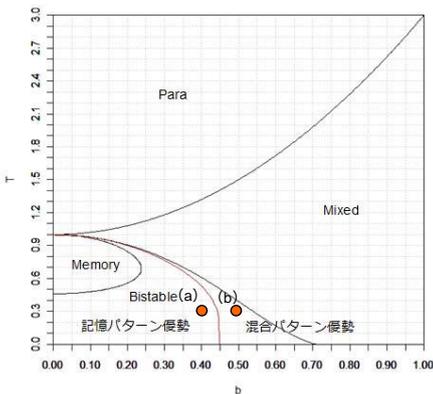


高次元空間上の様子



熱力学的エネルギーに関するイメージ図

混合パターン優勢の領域では、記憶パターンの想起は失敗し、混合パターンが想起される。この混合パターンは親パターン数と一致すると考えられ、その反転パターンも含め2個のパターンの出現確率が高い。



ピアソンドイバージェンス最大化に基づく 確率密度比推定のための次元削減手法

杉山 将, 原 聡, Paul von Buena, 鈴木 大慈, 金森 敬文, 川鍋 一晃
 東京工業大学 東京工業大学 Technical University of Berlin 東京大学 名古屋大学 Fraunhofer FIRST.IDA

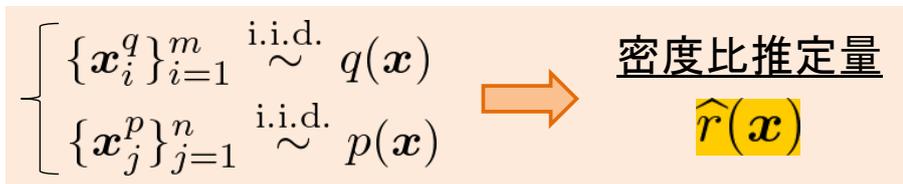
■ 確率密度比

$$r(\mathbf{x}) = \frac{q(\mathbf{x})}{p(\mathbf{x})}$$

応用例: 非定常環境適応,
外れ値検出, 特徴選択など

■ 確率密度比推定

KLIEP (NIPS'07, Sugiyama et al.)
uLSIF (NIPS'08, Kanamori et al.)



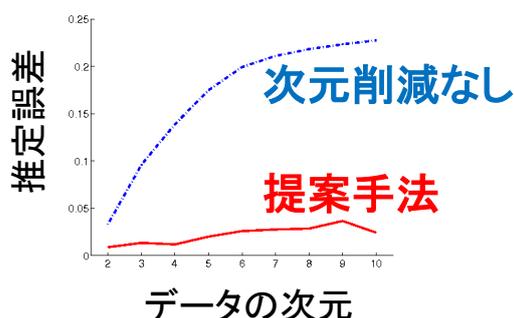
■ 目標: 高次元データの確率密度比の推定精度向上

提案手法

- 「次元削減 + 確率密度比推定」による精度向上
- 2つの密度関数が等しい部分空間は密度比に寄与しない
 ⇒ 2つの密度関数の差異が最大の部分空間を探索

$$\Rightarrow \frac{q(\mathbf{x})}{p(\mathbf{x})} = \frac{q(W\mathbf{x})}{p(W\mathbf{x})} \quad (W: \text{射影行列})$$

■ 人工データ実験結果

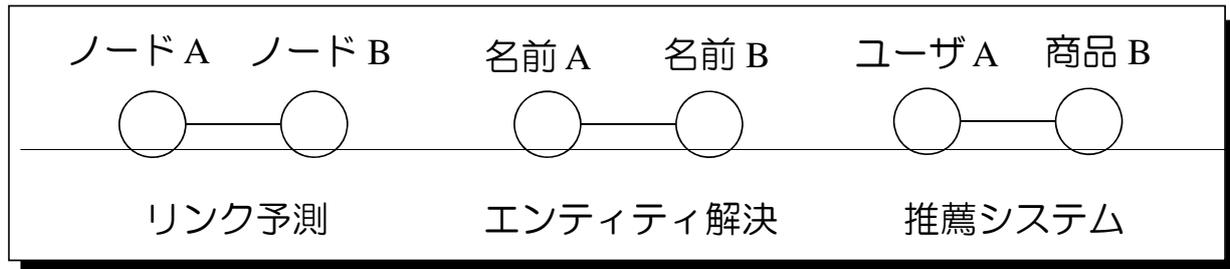


提案手法により
推定精度が向上

[74] 多様体正則化を用いたペアワイズ分類器の半教師付き学習

小山 聡(北大) / 鹿島 久嗣(東大)

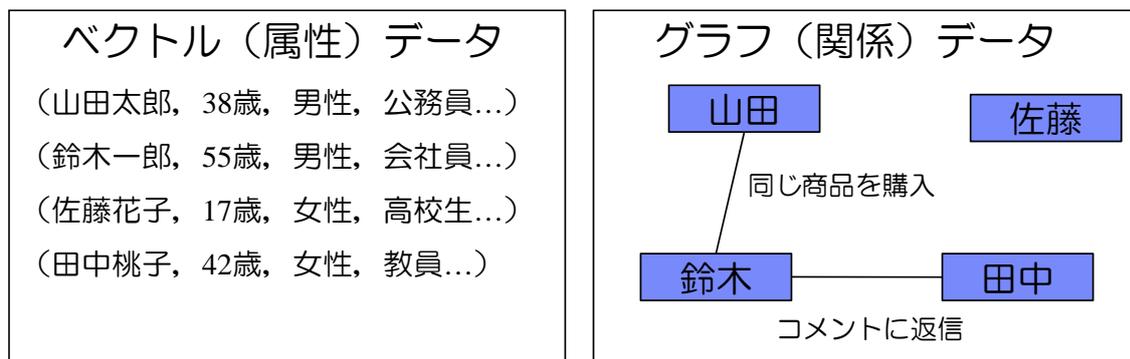
我々は、ペアワイズ分類問題



をうまく解くためにベクトル（属性）データとグラフ（関係）データを組み合わせる方法を考えます

「多様体正則化」によって
ペアワイズ分類における「Cold Startの問題」を解決します

- ペアワイズ分類には、ベクトル（属性）データとグラフ（関係）データの両方を用いた方が性能がよいことが知られています



- しかし、学習したモデルを新規データに適用する際には、グラフデータは利用できない場合が多くあります（=cold start 問題）
- 今回は、多様体正則化の枠組みを用いて、訓練時にのみグラフデータを利用する方法を提案します

No.81 候補系列抑制による高速系列データマイニング

水上 紘悠*, 岩沼 宏治**, 鍋島英知**

研究背景: 系列データマイニング

新聞記事系列データ

時系列



単語集合の系列データ

〈(台風 送検...) (洪水 初公判...) (判決 大雪...)... (台風...) (洪水...)...〉

〈(台風) (洪水)〉 有用頻出パターンとして抽出

研究目的

パターンの抽出をアプライアアルゴリズムに基づき行くと、段階的に**頻出な系列**を抽出する事ができる。しかし、膨大な数の**候補系列**(頻出の可能性を持つ系列)を生成してしまうという問題がある。



候補系列数を抑制する事で高速化を図る

* 山梨大学大学院医学工学総合教育部修士課程コンピュータ・メディア工学専攻

** 山梨大学大学院医学工学総合研究部

候補系列の抑制

主な手段として、ハッシュを用いて候補系列の抑制を行う。

概要:

step1. 頻出系列 L_k を求めると同時に、長さ $k+1$ の系列を枚挙し、ハッシュテーブルに格納する

step2. 長さ $k+1$ の候補系列生成時にテーブルの値を元に生成の判定を行う

実験結果

系列長	抑制無し	table_size				頻出系列
		256	4096	16384	262144	
2	38416	38416	13166	2234	565	483
3	10781	10781	2984	696	334	310
4	1810	1792	126	48	42	42
5	74	9	3	3	3	3
6	3	2	2	2	2	2
time [s]	4.14	3.83	2.29	1.57	1.42	

実験パラメータ

系列長5000, アイテム種類数1127, 1日1アイテム, 最小サポート0.1%

十分なサイズのハッシュテーブルを用意することで、頻出系列数と同等な値まで候補系列数を抑える事ができる。また、候補系列の抑制に伴い、実行時間に関しても向上がみとれる。

図: 候補系列数と実行時間

文献 Jong Soo Park, Ming-Syan Chen and Philip S. Yu:

An Effective Hash-Based Algorithm for Mining Association Rules.

In Proc. ACM SIGMOD Int. Conf. Management of Data, pp.175-186, 1995.

[P085]

A self-organizing state space approach to infer time-varying causalities between regulatory proteins

(自己組織化状態空間モデルを利用した転写調節タンパク質間時変因果の推定)

広瀬 修, 清水謙多郎

東京大学農学生命科学研究科
応用生命工学研究室

状態空間モデルによる転写調節タンパク質間の時変因果推定

モデル

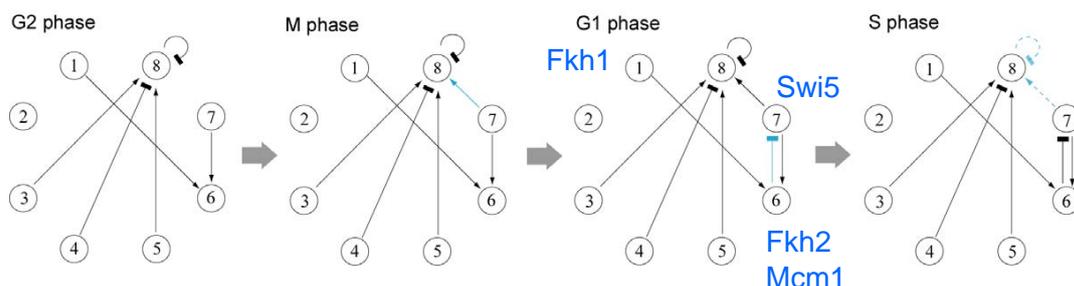
- 観測モデル $y_n = Hx_n + w_n$ $w_n \sim N(0, R)$
- システムモデル $x_n = F_n x_{n-1} + v_n$ $v_n \sim N(0, Q)$
- 初期状態 $x_0 \sim N(\mu_0, \Sigma_0)$

$y_n \in \mathbb{R}^p$: 遺伝子発現ベクトル, $x_n \in \mathbb{R}^k$: 状態ベクトル

罰則尤度

$$l(\theta) = \ln \sum_{l=1}^L p(y_1, \dots, y_N; \theta) - \lambda \sum_{n=1}^{N-1} \sum_{i,j} (f_{ij,n+1} - \eta f_{ij,n})^2$$

推定された転写調節タンパク質間の時変因果ネットワーク



- Swi5によるM/G1期の遺伝子群の発現の活性がFkh2の結合により抑制される. (Voth et al. 2007)
- Fkh1とFkh2はM期に必要なとされる遺伝子群の転写をMcm1と強調することで活性化する. (Kumar2000)

88 多タスク学習のための二次錘計画法

加藤 毅, 鹿島 久嗣, 杉山 将, 浅井 潔

お茶大

東大

東工大

東大, 産総研

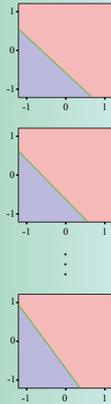
kato-tsuyoshi@aist.go.jp



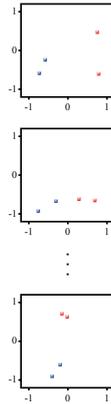
多タスク学習 (Multi-Task Learning)

関連するタスクを同時に学習して, 学習精度を向上させる

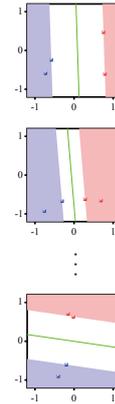
真の境界



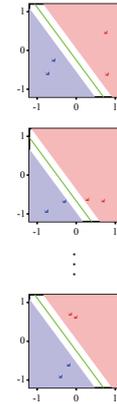
学習用例題



既存 SVM の
学習結果



多タスク学習の
結果



SVM をベースに二次錘計画問題として多タスク学習法を定式化



88 多タスク学習のための二次錘計画法

加藤 毅, 鹿島 久嗣, 杉山 将, 浅井 潔

お茶大

東大

東工大

東大, 産総研

kato-tsuyoshi@aist.go.jp



応用例

- 順序回帰
- 多クラス分類
- ネットワーク予測
- 協調的フィルタリング

参考文献

- Kato et al., Multi-Task Learning via Conic Programming, NIPS 2007.
- Kato et al., Conic Programming for Multi-Task Learning, TKDE, in press.



• 背景・目的

- 協調フィルタリングは評価履歴が一定数以上あるユーザ、コンテンツに対しては強力だが、新たなユーザ、コンテンツに対応できないという欠点をもつ
- 属性ベースフィルタリングはユーザやコンテンツの属性を利用することで新規なユーザ、コンテンツにも対応できるが、推薦精度が悪いという欠点を持つ
- ユーザ、コンテンツに属性情報がついている場合、それらの情報を最大限に利用することで両者の長所を併せ持つ推薦手法を実現したい

• アイデア

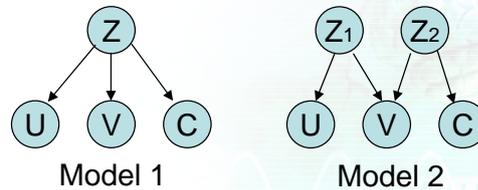
- 協調フィルタリングをソフトな Co-Clustering として実現する
- (事後的に)属性とクラス帰属度とを関係づける

• モデル

- Model1: pLSI (Aspect Model)
- Model2: Co-Clustering

• データ

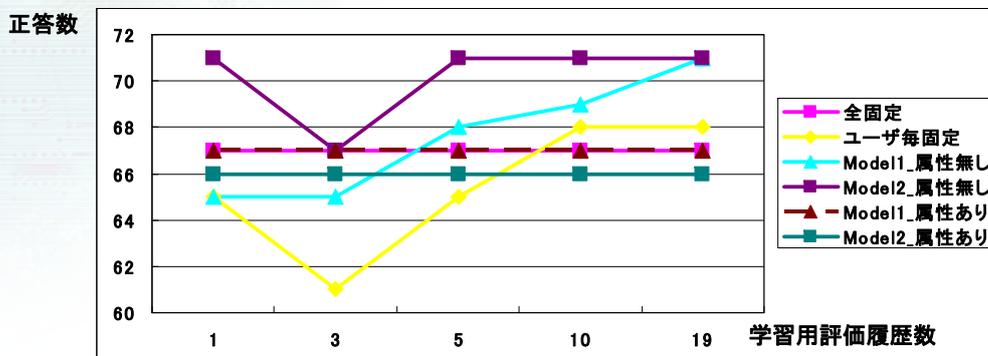
- AIST-KDDI研 映画嗜好調査
1616人 20作品/人の評価



結果と考察

• 結果の一部 (最尤学習、2値評価の正答率、潜在クラス数=2)

- 評価用ユーザ100人の評価履歴数を変えて学習(その他ユーザは19件)



• 考察と課題

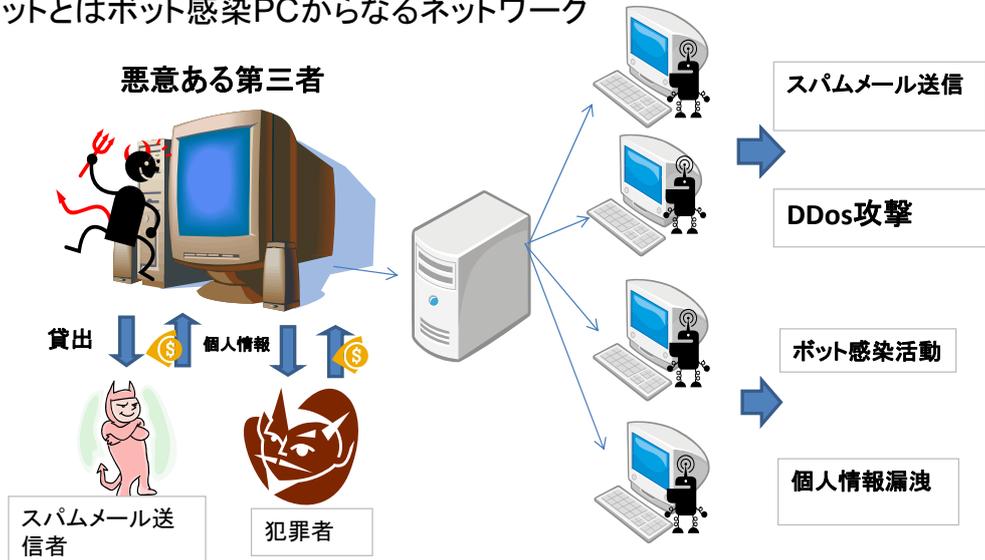
- 履歴が少ないときに属性を使うことで性能を改善できることを確認 (Model1)
- 履歴が多いときには属性を使うと性能が劣化する
- ベイズ学習化
 - 評価履歴が溜まると自然に属性ベースから履歴ベースに移行するように

P103 スパース構造学習によるボットネット検出の検討



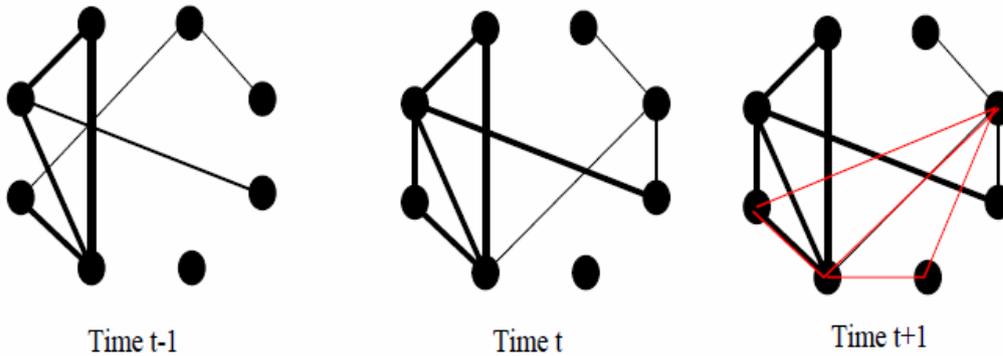
九州大学,九州先端技術研究所 村上慎太郎,濱崎浩輝,川喜田雅則,竹内純一
横浜国立大学 吉岡克成
情報通信研究機構 井上大介,衛藤将史,中尾康二

ボットネットとはボット感染PCからなるネットワーク



提案手法の概要

グラフの変化を検出



- ・ノードはソースホストを表す
- ・各ホストからセンサ(ダークネット)に届くパケット数の時系列を観測
- ・エッジの太さはノードの時系列間の相関の強さ(エッジがないものは相関なし)

バックグラウンドの通信が定常ならボットによる通信が行われた場合

ボットのペアに対応する共分散が一様に増加する

ただしデータが高次元小標本であること、ほとんどの通信が独立であることから

共分散行列の推定をスパース構造学習であるglasso(Friedman,2009)を用いて行う



P106 隠れマルコフカーネルマシンを用いた系列データの識別とその音素認識タスクへの適用

* 早稲田大学基幹理工学研究所

久保陽太郎*, 渡部晋治**, Erik McDermott**, 中村篤**, 小林哲則*

** NTTコミュニケーション科学基礎研究所

背景

1. 音声認識は発話の尤度を最大にするHMM状態系列を探索する問題
2. 非線形の識別を行なうため出力分布には混合ガウス分布が用いられる
3. 系列カーネルを用いたSVMが成果を上げている

問題

1. 系列カーネルを用いたSVMは汎化性能が高いがHMM状態のような概念を持たず探索アルゴリズムが適用できない
2. GMMを出力分布としたHMMは過学習しやすく局所解に落ちやすい

目的

HMMの出力分布関数をカーネル法によって拡張し、混合分布の利用を避けながら非線形の識別を実現

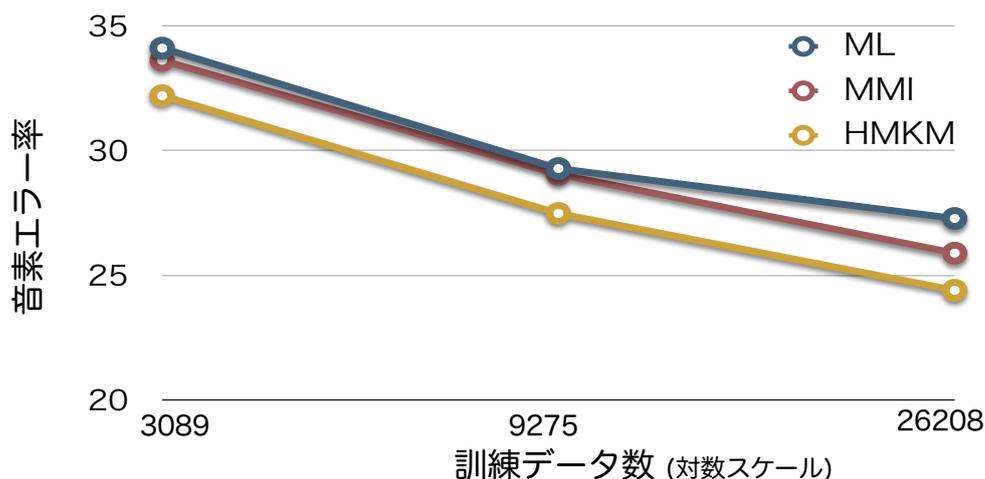
手法

1. HMMの出力分布として正規化されていない対数線形分布を利用
2. 識別学習を凸最適化問題に緩和して学習
3. 双対表現にカーネルトリックを適用し出力関数をカーネル化

実験

孤立音素識別タスクを行なった

比較対象： ML = 最尤推定によるHMM, MMI = 識別的基準で学習したHMM



考察

1. データが不足している時に過学習が避けられているのが確認できた
2. データが充分にある (MMIが充分な性能を持つ) 時にも性能向上
→ 局所解が避けられたため → 正則化による汎化性能向上

課題

連続音声認識への適用と評価



第12回 情報論的学習理論ワークショップポスターレビュー集

編集	第12回情報論的学習理論ワークショップ実行委員会
発行年月日	2009年10月19日
	ISBN 4-9902248-5-X

本ポスターレビュー集に掲載された資料の著作権は著者自身に帰属します。