Poster T-1

Decoding an Unknown Channel with Correlated Signals

Jian-Qin LIU^{*} ⁺ Teruya YAMANISHI[‡] Haruhiko NISHIMURA[¶] Hiroaki UMEHARA[†]

+ National Institute of Information and Communications Technology + Fukui University of Technology • ¶University of Hyogo * Corresponding author

Abstract With respect to an unknown channel that is defined by correlated signals of the input and the output, the signals of a sender T_x and a receiver R_x can be regarded as identical when the correlation value of these signals is above a threshold. In this report, we discuss the decoding issue on such a channel.

Keyword Decoding, Channel Model, Information Theory

The theoretical study on the issue of decoding an unknown channel, where we are focusing on the mathematical analysis of the signal processing aspect of the communication process under the condition of the estimated parameters. The formalized model of the object is illustrated in Fig. 1 where a sender T_x and a receiver R_x correspond to the input and the output, respectively.



Main Results

Lemma 1

If a correlation function $\psi(x,y)$ is maximized, it can be inferred that the corresponding mutual information I(x,y) is maximized under the condition of free of any channel delay.

Theorem 1

If correlation $\psi(x,y)$ is maximized, it can be inferred that the corresponding mutual information I(x,y) of the channel with the channel delay is maximized.

Conjecture 1

The minimum of the rate distortion is inferred from the optimization of the procedure for calculating the value of z subject to the minimum of the channel loss.

Conclusion

In order to study the information theoretic aspect of an unknown channel with correlated signals, the communication channel model is formalized under the condition of a nonlinear dynamics of the transmission process. Then, it is extended into a MIMO model that is equivalent to an interference channel where a Gaussian distribution is adopted for the decoding process. Considering the phenomenon of the synchronization (correlated signals) that is commonly observed in nature (e.g., signals in the brain) and engineering (e.g., the electric signals in circuits), the decoding issue (e.g., in the example illustrated in Fig. 1) is significant for systematically understanding the complexity mechanism of communication processes in information theory.

T-2 変分ベイズ学習の局所解

東京工業大学 中村文士•渡辺澄夫

モデル: p(x) = (1-a)N(x) + aN(x-b)

a,bの推定を変分ベイズ法で行う





^{T-3} 各成分の分散値が異なるランダム 行列の漸近固有値分布

新里 隆(秋田県立大)



GGMがクラスタ構造を有するように学習される?

本研究の概要

経験的事実の確認 GGMがクラスタ構造を有する ようには学習されない

■新たな正則化項の導入

- 既存クラスタGGM学習法よりも 簡便な手法の構築
- クラスタ数は自動的に決定









理論・最適化・応用等でアドバイスお願いします. キーワード: スパース, 低ランク学習, クラスタ, グラフ・ラプラシアン, グラフィカル・モデル

T-5:相対確率密度比 山田誠(東工大), 鈴木大慈(東大), 金森敬文 (名大), 八谷大岳, 杉山将 (東工大)

■確率密度関数の比(確率密度比):

$$r(\boldsymbol{x}) = rac{p(\boldsymbol{x})}{q(\boldsymbol{x})}$$

確率密度比は様々な統計的データ解析に役立つ! 非定常環境下での適応学習,外れ値検出, 二標本検定,次元削減,因果推論,パターン認識, 特徴選択,変化点検出,クラスタリング等

> Sugiyama et al. (NIPS 2007), Kanamori et al. (JMLR 2009), Suzuki & Sugiyama (AISTATS 2010), Yamada et al. (AISTATS 2011), Sugiyama et al (ICML 2011), Yamada et al. (AAAI 2011) etc.



本発表では理論解析&実験から相対確率密度比 が有用であることを報告!





T-7 Approximate Reduction from AUC Maximization to 1-norm Soft Margin Optimization 末廣大貴·畑埜晃平·瀧本英二(九州大学)



- 問題のサイズが p+n から pn に増大
- ▶ 既存の効率化手法はマージンに関する理論保証なし

提案手法

・最適化問題の p n 個の制約条件を p + n 個の制約条件に近似(計算量改善)
 ・マージンに理論保証有り

<u>実験結果(人エデータ)</u>

AUC 比較

計算時間比較 (sec.)

データ	Rank Boost	SoftRank Boost	ナイーブ	提案 手法
А	0.8979	0.7391	0.9125	<u>0.9940</u>
В	0.8857	0.9043	0.9136	<u>0.9173</u>
C	0.8727	0.8690	0.9043	<u>0.9007</u>

サンプル サイズ	ナイーブ	提案 手法
100	0.10	0.11
500	24.51	0.51
1000	256.78	0.86
1500	1353.00	1.76
3000	N/A	5.94

・AUCはナイーブ法に匹敵
・計算時間はナイーブ法より超高速

T08: Kernel and Feature Search in Kernel PCA

Alam Md. Ashad and Kenji Fukumizu

The Graduate University for Advanced Studies and The Institute of Statistical Mathematics, Japan

Motivation of Our Study:

- The principle of kernel methods is to seek a new space by using a desired kernel.
- Appropriate method for kernel and feature choice in kernel PCA have not yet been established.
- We are not able to choose them based on a performance with the norms of the feature spaces.
- We have proposed a method for choosing optimal kernel and the number of features through the leave-one-out cross-validation based on the performance of pre-images.

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

Experimental Results: synthetic data as well as 4 real world problems

	Wine		Diabetes			
$s ackslash \ell$	2	3	2	3	4	5
0.1	4.29775	4.25167	2.24101	2.26142	2.40629	2.39790
	<u>0</u>	(3.33)	(37.78)	(22.22)	(28.89)	(26.67)
0.25	<u>3.98027</u>	4.32378	1.99628	2.05311	<u>1.97823</u>	2.12594
	(<u>1.67</u>)	(3.33)	(33.33)	(22.22)	(<u>13.33</u>)	(26.67)
0.5	4.10606	4.07721	2.23193	2.09396	2.16748	2.08503
	(1.67)	(5.0)	(20.00)	(28.89)	(22.22)	(26.67)

Table: Wine and Diabetes data sets: leave-one-out cross-validation error and classification errors for test data (in bracket) using RBF kernel.

For classification data sets our method chooses the kernel and the number of features that give the best classification errors in most cases, we can confirm that our method successfully finds the parameters that give a reasonable low-dimensional representation in kernel PCA, **T9:** サンプル間の相関を考慮した主成分分析と カーネル関数による共分散行列の推定

林浩平,川喜田雅則,池田和司





有

主成分分析 (PCA) を拡張し,相関のあるデータに適し たモデルを提案

無

サンプル間の共分散をノンパラメトリックに推定

実験結果 PCAに比べパラメータの推定誤差が減少



T-10状態遷移の推定に基づく 能動的価値関数推定法





T-11 神経細胞の膜電位時系列から時間変動する 入力信号と隠れ変数のダイナミクスを推定する 小林 亮太¹, Lansky Petr², 篠本 滋³ 1: 立命館大学, 2: チェコ科学アカデミー, 3: 京都大学 kobayashi@cns.ci.ritsumei.ac.jp

- 神経細胞:
 - 高次元の状態変数で記述される.
 - 時間変動するシナプス入力信号を受けている.
 - 1変数 (膜電位)しか測定できない.
- 問題: 膜電位時系列から入力信号と測定できない状態変数のダイナミクスを推定する.
 - 入力信号,状態変数ともに時間変動する.
 推定パラメータ数 > データ数 (不良設定問題)

T-11 神経細胞の膜電位時系列から時間変動する 入力信号と隠れ変数のダイナミクスを推定する

- 推定手法:
 - ガウシアンプロセス型事前分布
 - 経験ベイズ法
- 推定例



T12: カーネル法に基づく行列あるいはテンソル補完

林浩平,竹之内高志,冨岡亮太,鹿島久嗣

目的:部分的に欠損した行列(テンソル)の補完 • 応用例:推薦システム



方法:観測そのものから類似度を計算する,新しい カーネル法の提案

- 効率的なアルゴリズム: O(RC(R+C))
- 真の行列 (テンソル) がフルランクでも予測可



実験結果

真のデータが低ランクでない場合,比較手法を上回る 予測性能を達成



- 報酬 ■近年、税金取り立てシステム、対話システム、 コンピュータ操作支援システムなどへの応用も盛ん!
 - ■本研究では、<mark>修正Newton法</mark>を用いた収束の速い 方策勾配法を提案する



T-14 最尤法における 潜在変数推定の理論解析 An Accuracy Analysis of Latent Variable Estimation with the Maximum Likelihood Estimator

山崎 啓介 (東京工業大学)

階層型モデルの潜在変数:推定誤差は未解明

潜在変数推定誤差の漸近形を得た
推定法:最尤法
誤差関数:潜在変数分布間の

$$KLダイバージェンス$$

 $D(n) = \frac{1}{2n} Tr[K(w^*)I_x^{-1}(w^*)]$
 $K(w) = I_{XY}(w) + I_X(w) - J(w) - J^T(w)$

離散DC計画問題へのプリズム法とその応用 河原吉伸, 鷲尾隆(大阪大学/科学技術振興機構)

DC = <u>D</u>ifference of two <u>C</u>onvex functions



離散DC計画問題へのプリズム法とその応用 河原吉伸, 鷲尾隆(大阪大学/科学技術振興機構)

DC = <u>D</u>ifference of two <u>C</u>onvex functions

DC計画問題(連続関数):



凸関数



- 本発表の内容(<u>NIPS'11</u>へも採録済み):
 - 離散DC計画問題への分枝限定法に基づく厳密解法を提案
 - Lovasz拡張を通した凸性との関係を用いて、問題の構造を利用した分枝限定法を構成.
 - 機械学習で扱われる具体的な問題例を見ると共に、いくつかの数値例を紹介する.

T-16 ガウス性に基づく マルチカーネル学習 日野英逸, 村田昇 (早稲田大学)

概要

特徴空間におけるデータ分布の最適化に基づくMultiple Kernel Learning(MKL)の新しい枠組みを提案. 特徴空間でFDAによる判別を行うとすると, 各クラスでデータが正規分布に従い,共分散構造が等し い状況が理想的 ← MKLによってデータ分布を動かす.

利点: カーネル結合パラメタに関する最適化のみでよい.

正規性の尺度として,経験 特性関数を利用. 特性関数は内積のみで記述 できるので,特徴空間にお ける分布をカーネル関数の みで評価可能.

正規性の評価は特性関数の squared moduloのみに着 目すれば良いという性質を 利用し, カーネル結合係数 に関する凸最適化問題とし てMKLを定式化.



T-I7 カラーチャンネル間の相関 を考慮したGaussian FoEモデル 東北大学大学院 「報科学研究科 村山 竜太,安田 宗樹,和泉 勇治,田中 和之 {ryuta,muneki,wai,kazu}@smapip.is.tohoku.ac.jp

 機械学習を通してカラーチャンネル間の相関を考慮に入れた 画像の事前確率を設計する。

• 関連研究:

S. Roth et alによるFoEモデル

 本講演ではRGBカラーチャンネルのピクセル間の確率的な 相互作用を取り込んだモデルを提案する。







モデル学習の結果、 カラーチャンネル間には 確率的な相互作用がある

画像補修フィルタに応用





	従来手法	提案手法
MSE	162.63	156.11
PSNR	26.05	26.21



事象間の共起関係を考慮したクラスタリングによる 燃料電池の損傷パターン抽出

稲場 大樹*1 福井 健一*2 佐藤 一永*3 水崎 純一郎*4 沼尾正行*2

*1 大阪大学大学院情報科学研究科 *2 大阪大学産業科学研究所 *3 東北大学工学研究科 *4 東北大学多元物質科学研究所







★ ℓ1 アルゴリズムの性能 (Donoho-Tanner転移)



統計力学(レプリカ法)で転移線の評価が可能

レプリカ法の解析ではレプリカ対称性を仮定 しかしレプリカ対称性が破れている場合も有る

 結論
 *レプリカ対称性が破れていても、対称性を 仮定し得られた転移線は直ちに否定されない
 *レプリカ対称性が破れる線は
 求解困難性が生じる別の転移になっていそう



T20: Generalization of Matrix Factorization for Robust PCA

中島 伸一(ニコン)、杉山 将(東工大)、デリン ババカン(イリノイ大)

行列分解モデル: $Y = BA^{\top} + \mathcal{E}$

分解によるモデル起因正則化 (Model-induced regularization)



変分ベイズ解はスパース [Nakajima&Sugiyama:JMLR2011]


Nikon

一般化行列分解 (Generalized Matrix Factorization)

$$Y = \sum_{s=1}^{S} G(\{B^{(k,s)}A^{(k,s)\top}\}_{k=1}^{K^{(s)}}; \mathcal{X}^{(s)}) + \mathcal{E}$$

$$U = \begin{pmatrix} \underbrace{U_{1,1} \ U_{1,2} \ U_{1,3} \ U_{1,4}} \\ \underbrace{U_{2,1} \ U_{2,2} \ U_{2,3} \ U_{2,4}} \\ \underbrace{U_{3,1} \ U_{3,2}} \\ \underbrace{U_{4,1} \ U_{4,2}} \\ \underbrace{U_{4,3} \ U_{4,3}} \\ \underbrace{U_{4,3} \ U_{4,4}} \end{pmatrix} \leftarrow \begin{pmatrix} U^{(1)} = \underbrace{(U_{1,1} \ U_{1,2} \ U_{1,3} \ U_{1,4}} \\ \underbrace{U_{2,1} \ U_{2,2} \ U_{2,3} \ U_{2,4}} \\ \underbrace{U^{(2)} = \underbrace{(U_{2,1} \ U_{2,2} \\ U_{2,3} \ U_{2,4}} \\ \underbrace{U^{(3)} = \underbrace{(U_{3,1} \ U_{3,2} \ U_{4,1} \ U_{4,4}} \\ \underbrace{U^{(4)} = \underbrace{(U_{3,3} \ U_{4,3}} \\ U^{(6)} = \underbrace{(U_{3,4} \ U^{(5)} = B^{(4)} A^{(4)\top} \\ U^{(6)} = \underbrace{(U_{4,2} \ U^{(6)} = B^{(6)} A^{(6)\top} \\ \underbrace{U^{(6)} = \underbrace{(U_{4,2} \ U^{(6)} = B^{(6)} A^{(6)\top} \\ \underbrace{U^{(6)} = \underbrace{(U_{4,2} \ U^{(6)} = B^{(6)} A^{(6)\top} \\ \underbrace{U^{(6)} = \underbrace{U^{(6)} = B^{(6)} A^{(6)\top} \\ \underbrace{U^{(6)} = B^{(6)} A$$

種々のスパース性を誘起する行列分解モデルの統一的枠組み (ロバストPCA [Candes+09]の拡張)

変分ベイズ行列分解の解析解 [Nakajima+:NIPS2011]





Time(sec)

T-21 再正規化最尤符号を用いた クラスタリング構造変化の検出

クラスター数が変化する時刻を検出する

東京大学

<u>平井聡</u>、山西健司



ラスター数が変化する時刻を検出したい!

手法の特徴

- 動的モデル選択を逐次的に近似
 動的モデル選択の手法を逐次的なデータに適用
- ▶ RNML符号長を評価規準として用いる
 - クラスタリングをガウス混合分布でモデル化
 - 。クラスター数を高い確率で当てることができる



T-22 協調フィルタリングによる商品販売量の予測

Tor Andre MYRVOLL 松井知子 SINTEF, Norway

統数研

- 仮定:ユーザの各商品の購買パターン~商品に対する嗜好
- 日本のあるスパーマーケットのPOSデータを利用
 - ユーザID: *p*_u
 - 商品ID: q_i
 - 日付(年、月): s_v , o_m
 - 購買数: r_{ui}

$$O(Q,P) = \frac{1}{2} \sum_{(u,i)\in\Omega} \left(r_{ui} - q_i^T p_u \right)^2 + \lambda \left(\|q_i\|^2 + \|p_u\|^2 \right) \quad \text{[Koren et al., 2009]}$$

$$O(Q,P) = \frac{1}{2} \sum_{(u,i,m,y)\in\Omega} \left| r_{ui} - q_i^T p_u - q_i^T o_m - q_i^T s_y - p_u^T o_m - p_u^T s_y - o_m^T s_y \right| + \lambda \left(\left\| q_i \right\|^2 + \left\| p_u \right\|^2 + \left\| o_m \right\|^2 + \left\| s_y \right\|^2 \right)$$

表1: テストセットに対する誤り || r - r' || の平均

Estimator	Mean test error
r' _{ui}	1.21598
r' _{uim}	1.15113
r' _{uimy}	1.13248

疎でないマルチプルカーネル学習の速い収束レート および最適な正則化に関して IBIS 2011: Poster ID T-23 | 鈴木 大慈 (東京大学) 実験:デンスな正則化→良い性能 Multiple Kernel Learning (MKL) $L\left(\sum_{1}^{M} f_{m}\right) + C \|(f_{m})_{m=1}^{M}\|_{\psi}^{2}$ Reuters (acq) 0.62 0.6 L1 0.58 U.58 USW WH 0.56 L₂ は 様々な正則化 0.54 0.52 **L**4/3 3000 5000 1000 2000 4000 6000 [Cortes: ICML2009 Invited Talk] $\|(f_m)_{m=1}^M\|_{\psi} = -$ **Elastic-net** なぜ? $\lambda \|f\|_{\ell_1} + (1-\lambda) \|f\|_{\ell_2}$ $\left\{ \sum_{j=1}^{M'} \left(\sum_{k=1}^{M_j} \|f_{j,k}\|_{\mathcal{H}_{j,k}}^p \right)^{\frac{q}{p}} \right\}^{\frac{1}{q}}$ 統一的な理論 <u>任意の</u>ノルム型正則化に適用可能! **VSKL**

疎でないマルチプルカーネル学習の速い収束レートおよび最適な正則化に関して IBIS 2011: Poster ID T-23 鈴木 大慈(東京大学)

各 RKHSの 複 雑 さ





【NMFの一般化】

【部分空間法の観点からの判別問題への応用】

部分空間法に基づく判別の概要

- 1. 観測データからクラス毎に部分空間を作成する
- 2. 未知データに対して1番近い部分空間を構成すのに用いたクラスを未知データのクラスとする



(C. Genest, 1987)





洁果



D-103 ROC曲線を考慮した記憶容量を超えた場合の パーセプトロンの学習

桑原昭之, 筑波大学大学院システム情報工学研究科 kuwa@bcl.esys.tsukuba.ac.jp





・(学習域内の)サンノルの相対: 頻度差を学習に反映

D-104文書情報を考慮した無限関係 モデルによるネットワーク推定

鈴木 知彦 +,井田 安俊 +,松本 隆 + (+早大) suzuki08@matsumoto.eb.waseda.ac.jp



- ネットワーク上で遣り取りされる文書の 生成モデルを提案
- クラスタ構造と文書
 モデルを同時推定



D-105 半教師あり学習を用いた ノンパラメトリックトピックモデルによる web文書の評判分析

井田 安俊, 鈴木 知彦, 松本 隆 (早稲田大学)

ida11@matsumoto.eb.waseda.ac.jp

目的:web上の製品レビューなどの文書に対して点数を予測.

アプローチ:

- 点数付き,もしくはラベル付き文書に対するトピックモデルの提案.
- ノンパラベイズ:文書のトピック数は無限.
- 単語の潜在変数をcollapsed Gibbs samplerで推定.
- 推定したトピックを使って最尤推定で点数を線形回帰.
- 半教師あり学習.



特徴の出現頻度に応じたL1正則化を実現するDual-Averaging D-106 O大岩秀和 松島慎 中川裕志 (東京大学)

- Dual-Averaging Method (DA) [Nesterov, 09]
 - 凸最適化問題の最適解を求めるオンライン学習手法

$$\begin{split} \min_{\mathbf{w}} E_{\mathbf{z}} f(\mathbf{w}, \mathbf{z}) &\approx \min_{\mathbf{w}} \sum_{t} f(\mathbf{w}, \mathbf{z}_{t}) \\ \mathbf{w} \in W \quad \text{if} : W \times Z \to \Re \quad \text{if} \in \mathbb{R} \\ \mathbf{w} \in W \quad \text{if} \\ \mathbf{w}$$

- 1データごとにパラメータを更新
 - 劣勾配空間と重み空間を分けて最適化

$$\mathbf{g}_t \in \frac{\partial f}{\partial \mathbf{w}_t}(\mathbf{w}_t, \mathbf{z}_t) \subset G \quad \Longrightarrow \quad \mathbf{w}_{t+1} = \operatorname*{arg\,min}_{\mathbf{w} \in W} \left\{ \frac{1}{t} \sum_{\tau=1}^t \langle \mathbf{g}_\tau, \mathbf{w} \rangle + \frac{\beta_t}{t} h(\mathbf{w}) \right\}$$

- Regularized Dual-Averaging (RDA) [Xiao, 2010]
 - L1正則化項等の正則化項を導入可能
 - 低頻度の特徴が零化されやすい 有効な特徴も、低頻度だと零化されやすい



提案手法

- 低頻度かつ有用な特徴が零化されにくいL1正則化を実現
 - 動的に、L1正則化による零化の効果を調節
 - 前処理を必要としない



- Regret上限 $O(\sqrt{T})$ を証明
- 実データを用いた実験により、既存手法を上回る精度を確認



[D107] **予測分布を考慮したマルチタスク学習** 成田敦博 佐藤一誠 中川裕志

マルチタスク学習:

□ 複数の学習器のあいだで情報を共有する

- タスク間のパラメータを近づける $\|w_1 w_2\|^2$ • 一方、素性はそれぞれ頻度やスケールが違う
- 負の転移
 - 精度が悪化することもある
 - タスク間の"関連度"も上手く決めたい

パラメータの予測分布を用いた オンラインマルチタスク学習 Confidence Weightedアルゴリズムの拡張
 パラメータの分布を逐次的に学習する分類器



- 逐次的に予測分布を学習
- タスク間の関連も動的に推定

 $a_{1t}^{i+1} \propto a_{1t}^{i} \exp\left(-\theta \cdot \mathbb{D}\left(\mathcal{N}(\boldsymbol{\mu}^{t}, \boldsymbol{\Sigma}^{t}) \| \mathcal{N}(\boldsymbol{\mu}^{1}, \boldsymbol{\Sigma}^{1})\right)\right)$

• 転移学習にも利用できる



ベータダイバージェンス最小化に基づく最適フィルタの

提案



ベータダイバージェンスとH。フィルタの関係

D-109 情報理論的規準に基づく新しい バースト検知アルゴリズムの クラスについて 金澤 宏紀*山西健司**東京大学院情報理工学系研究科 {hiroki_kanazawa, yamanishi}@mist.i.u-tokyo.ac.jp

✓ バースト検知基準 [Kleinberg 2003]
 バースト:時系列情報の(急な)変化
 例)メールの受信頻度の増加
 → どこをバーストとしてよいかわからない

受信インターバル $x^n = x_1 \dots x_n$, x^n は $f(x_i; \theta_i) = \theta_i \exp(-\theta_i x_i)$ に従うと仮定

 Θ を離散化して動的計画法で解き θ_{i_t} の大きい区間をバースト区間とする 規準: $\sum_{t=1}^{n} -\log f(x_t; \theta_{i_t}) + \sum_{t=1}^{n} \tau(i_t|i_{t-1})$ \rightarrow minimize w.r.t. (i_1, \dots, i_n)





問題点: 指数分布に限定的 離散化法がアドホック ペナルティ項(τ(・|・))の理論的意味づけがない

今回の成果: 一般のパラメトリックな分布に対する変化点検出 情報幾何における等距離則に基づく最適離散化法 情報論的規準に基づくペナルティ項の設定

これを起点にしてバースト検知アルゴリズムを構成 複数の理論的バリエーションを考えることができるので, 数値実験にてそれらを比較検討する

数值実験:

- → いくつかのパラメトリックな分布に対して適用する
- → 変化点(バースト含)が検知できることを確かめる

D110: 動的方策計画法によるモジュール学習制御

植野 剛1, 河原 吉伸2, 鷲尾 隆2

¹科学研究機構 湊離散構造処理系プロジェクト ²大阪大学産業科学研究所

<u>マルコフ決定過程 (MDP)</u>

 $p(\bar{x}, \bar{u}|x_0) = \pi(u_0|x_0) \prod_{i=0}^{n-1} p(x_{i+1}|x_i, u_i) \pi(u_i|x_i)$ 状態: $x \in \mathcal{X}$, 行動 $u \in \mathcal{U}$, 状態遷移確率: p(x'|x, u)報酬: $\ell : \mathcal{X} \times \mathcal{U} \mapsto \mathbb{R}$, 方策: $\pi : \mathcal{X} \mapsto \mathcal{U}$

*MDPは非線形、非定常性を示すとする

価値関数
$$J_{\pi}(x)$$
を最大とする方策 π^* を見つける $\pi^* = \operatorname*{argmax}_{\pi} J_{\pi}(x)$
 π \mathcal{L}

D110: 動的方策計画法によるモジュール学習制御

<u>後数モデル同定型強化学習(Multiple model-based reinforcement learning)</u>



予測器により**どの部分空間に属しているか予測**し、予測の確信 度に応じて各モジュールの出力を重み付け足し合わせる

(問題点) 方策学習則に対する妥当性が無く収束の保証はない



D-111 ポテンシャルガイディング関数を用いた 条件付き 状態遷移軌跡分布のモンテカルロ計算

松田 衆治, 伊庭幸人, 鷲尾隆(大阪大学, 統計数理研究所)

特定の始点から終点への状態遷移軌跡分布の導出
 モンテカルロ計算によるより効率的な導出



先行研究:Lattice Simulations of Polymer Growth [Robert 2008]
 選択的に状態を遷移させる事による効率的な状態遷移軌跡分布の導出

- 先行研究 [Robert 2008]における問題点
 - 選択的に遷移させる事で<mark>歪んだ結果</mark>になる.



- 本研究の目的
 - ポテンシャル性を持つガイディング関数により、

分布を歪める事なく選択的に遷移させる.

- これにより, 正しい分布を保持しつつ, 効率化を図れる.



D-112 イベント系列データからの生活コンテキスト推定のための 確率的生成モデルの検討

黒川茂莉,村松茂樹,横山浩之(KDDI研究所) 吉井和佳,麻生英樹(産業技術総合研究所) mo-kurokawa@kddilabs.jp

目的:携帯電話の通信で利用した基地局の位置情報に基づいて、 利用者の日々の生活コンテキストを推定し、 生活コンテキストに応じた情報提供サービスを行う

課題:通信時の基地局位置情報のような、取得間隔が一定ではなく、 空間的にも粒度が粗い位置情報を、時間的・空間的に補間し、 日常/非日常の生活コンテキストを推定する

アプローチ:日常的な行動は周期的に繰り返されるという仮説をもとに、 ガウシアンプロセス(1日周期の周期カーネル)を用いて基地局位置情報の 時系列データを補間し、時系列データを日常行動部分と非日常行動部分 にふるい分けする





実験結果例





National Institute of Advanced Industrial Science and Technology **AIST** 2

マルチタイプ混合メンバーシップ・ブロックモデル D-113 を用いた情報推薦 目的 横峯 樹(神戸大学),江口 浩二(神戸大学) ・異種ネットワークのための混合メンバーシップブロックモデルの 提案 ・提案モデルによって、補助情報を用いた協調フィルタリングを実現 アプローチ ・ベイズ的アプローチによるネットワークモデルとして知られる混合 メンバーシップブロックモデルをリンクのタイプを扱える形に拡張 ・提案モデルでの情報推薦の精度を多角的に評価





t = 0, k = 1, h = 2のとき $0.3 \times 0.5 \times 0.4 = 0.06$

D-114 Symmetric Correspondence Topic Models による多言語トピック抽出

福増康佑(神戸大),松浦愛美(東大),江口浩二(神戸大)

研究内容

言語を横断したトピックの抽出に適した新たなトピックモデル(SymCorrLDA)の提案

多 言 語 デ ー タ 向 け に CorrLDA[†]を 拡張

- ▶ 元々は画像とそのアノテーション向けに提案されたトピックモデル
- ➢ CorrLDAを多言語データで使う場合、あらかじめ定めた主軸となる言語のトピックを推定し、それに基づき非主軸言語のトピックを推定する ⇒SymCorrLDAでは文書ごとに主軸言語を自動で調整するように改良



実験·比較

◆ 日英のWikipediaデータにおいて、提案手法を既存手法および多言語 データに適用可能な他手法と比較

単語毎テストセット対数尤度

	CI-LDA	SwitchLDA	CorrLDA (日本語主軸)	CorrLDA (英語主軸)	LDA	CorrLDA -random	SymCorrLDA
日本語	-8.14	-8.14	-7.46	-7.78	-8.13	-7.51	-7.43
英語	-8.64	-8.64	-8.40	-8.20	-8.63	-8.25	-8.17





対訳発見成功率





—:提案手法				
	両方の実験で			
	高いパフォー			
	マンスを運成			
	主軸言語を争 前に決定する			
	必要なし			


カーネルむかしばなし @ IBIS2011

カーネル 「2次曲線のあてはめは, (*x², xy, y², x, y, 1*)という特徴ベクト ルを使ったPCAで,小さい固有値に対応する固有ベクトルを求めればいい んですよね」

ティホノフ 「そのとおりじゃ」

- カーネル 「それじゃあ,カーネルPCAを使えば無限次元の究極の曲線あて はめができますね!」
- ティホノフ 「うーむ,残念ながらその方法にはリプレゼンターの力が働かな いのじゃ」
- カーネル 「えええっっっ!!!! ではどうしたらいいのですか?」

ティホノフ 「D-115 のポスターに急ぐのじゃ. ただし,別に答が得られ るわけではないがのう. おまえのアドバイスが役に立つに違いないぞ」



ガウス過程によるベイズ化

従来: ラプラシアン $\mathbf{f} = \left(L + \frac{1}{2\sigma^2} I \right)^{-1} \mathbf{y}$ (全てのラベルなし)² ラベル伝搬法 $\mathbf{f} = \left(L + \frac{1}{2\sigma^2} I \right)^{-1} \mathbf{y}$ サイズの行列. 大きい&重い! ガウス過程 $\hat{\mathbf{w}} = \arg \max p(\mathbf{w} | X, \mathbf{y})$ MAP推定 分類対象でないラベルなし 提案:ベイズ推定 データをここに詰め込める $p(\mathbf{y}_* | X_*, X, \mathbf{y}) = \int p(\mathbf{y}_* | X_*, \mathbf{w}) p(\mathbf{w} | X, \mathbf{y}) d\mathbf{w}$ 結果: 計算量削減&ベイズ化による判別性能向上

D-117 階層的発話生成モデルを用いた話者クラスタリングのための フルベイズモデル推定手法の比較

俵 直弘(早大), 小川 哲司(早大), 渡部 晋治(NTT), 小林 哲則(早大)



- 音声データを用いた**話者クラスタリング**に
- 適した**フルベイズモデル推定**手法の開発



- 話者の数が不明
- 話者内変動
 - 特徴量の分布が複雑で単峰の
 ガウス分布で表現しきれない

• 話者間変動

- 話者毎に発話数のばらつきがある



アプローチ:

Utterance-oriented DPMM (UO-DPMM)

● <u>発話生成モデル</u>

- 各話者の発話が話者固有のガウス分布から 独立に生成されると仮定
- 話者分布を混合要素とする混合分布として
 定式化することで話者内・話者間変動をモデル化

● <u>ノンパラメトリックベイズ推定</u>

- 話者数推定とクラスタリングを同時に行う

検討内容

- UO-DPMM
- 話者分布の混合分布化

● 最適なモデル推定手法の検討 (変分ベイズ法・サンプリング法)



D-118 マルチモーダル情報を用いた トピックモデルによる映像分類

梶野 洸(東大), 木村 昭悟(NTT), 石黒 勝彦(NTT)

- ▶ Web動画のカテゴリ分類の実現
 - ▶ 手法: <u>映像</u>と<u>音響</u>の両者を用いるトピックモデル



D-118 マルチモーダル情報を用いた トピックモデルによる映像分類

▶ 実験結果

- カテゴリによって結果が異なる
- ▶ 映像と音声が合成されていることがある
- ▶ 映像と音声との相関の有無を考慮した手法を提案



1004 - Airplane Flying



1105 - Singing



D-119 ループありネットワークの混合 ループなしネットワークへの変 換による推論・学習

前田 新一,石井 信 京大 情報学

ichi@sys.i.kyoto-u.ac.jp, ishii@i.kyoto-u.ac.jp

関連キーワード

Tree-reweighted algorithm



D-120. アクティブ計測とパッシブ計測を用いた ネットワークトモグラフィ 宮本 敦史, 渡辺 一帆, 池田 和司(奈良先端大)

ネットワークトモグラフィとは、

パスのパケットロス率からリンクのパケットロス率を求める 逆問題である(通常はアクティブ計測のみ)

• 本研究ではアクティブ計測とパッシブ計測の2つを用いる





- 計測精度を落とさず、計測コストを減少させられることを示す
- 複数の計測法を用いたネットワークトモグラフィ問題を定式化
- 推定アルゴリズムの導出

参考文献

- R. Caceres, et al., "Multicast-based inference of network internal loss characteristics", IEEE Trans. Inf. Theory, 1999.
- M. H. Firooz et al., "Network Tomography via Compressed Sensing" IEEE GLOBECOM, 2010.



D-121 混合整数計画による正則化 0/1 損失最小化

加藤 毅



混合整数計画による定式化1

 $\begin{array}{ll} \min & \|\boldsymbol{e}\|_{1} + C^{-1} \|\boldsymbol{w}\|_{\infty}, \\ \text{wrt} & \boldsymbol{e} \in \{0,1\}^{\ell}, \boldsymbol{w} \in \mathbb{R}^{d}, b \in \mathbb{R}, \\ \text{subj to} & \forall i \in \mathbb{N}_{\ell}: \quad y_{i}(\boldsymbol{w}^{\top}\boldsymbol{x}_{i} + b) \geq 1 - Me_{i}. \\ \text{where } C \in \mathbb{R}_{++} \text{ is constant.} \end{array}$

混合整数計画による定式化2

$$\begin{array}{ll} \min & \|\boldsymbol{e}\|_{1}, \\ \text{wrt} & \boldsymbol{e} \in \{0,1\}^{\ell}, \boldsymbol{w} \in \mathbb{R}^{d}, b \in \mathbb{R}, \\ \text{subj to} & \forall i \in \mathbb{N}_{\ell}: \quad y_{i}(\boldsymbol{w}^{\top}\boldsymbol{x}_{i}+b) \geq 1 - Me_{i}, \\ & \|\boldsymbol{w}\|_{\infty} \leq \lambda \\ \text{where } \lambda \in \mathbb{R}_{++} \text{ is constant.} \end{array}$$



D-122. ラージマージンなアンカーグラフハッシング

得居誠也⁺, 佐藤一誠⁺, 中川裕志^{+ +}東京大学院 情報理工学系研究科 ⁺東京大学 情報基盤センター

近傍探索のためのハッシング



少数のアンカーを経由するような類似度 [Liu+, 10] を用い ることで, 上のハッシュ値最適化問題の解を高速に求めるこ とができる [Liu+, 11]

2 つの問題点を最適化問題の変形によって改善

1. 連続緩和による 境界でのエラー



➡ 二値化ロスを導入してハッシュ関数をラージマージンに 変形することで、広めな近傍での精度を向上

2. 直交制約の非効率性



→ 組み合わせ直交制約を導入することで少ないビット数で も高い近傍精度を達成