

平成22年 5月24日現在

研究種目：特定領域研究

研究期間：2006～2009

課題番号：18079013

研究課題名（和文） 確率推論における近似解法の理論的解析および情報工学への応用

研究課題名（英文） Theoretical analysis and information engineering application of approximate inference methods

研究代表者

池田 思朗（ IKEDA SHIRO ）

統計数理研究所・数理・推論研究系・准教授

研究者番号：30336101

研究成果の概要（和文）：確率推論は情報工学，統計学などの様々な問題において基礎となる。近年，様々な分野で大規模な問題を扱うことが増え，確率推論を厳密に行うのではなく，近似的に行うことが増えている。本研究ではそのような近似手法に関わる数理的基盤を研究した。研究成果を下に衛星デジタル放送の受信結果を向上させるアルゴリズムを構築した。また，領域内の交流から神経細胞通信路モデルの研究，脳における運動計画モデル，多クラス判別問題についての研究を行った。

研究成果の概要（英文）：Stochastic inference is important in many problems in the fields such as information engineering and statistics. Recently, more attention is paid for approximate inference methods because exact inference becomes intractable for large size problems. In this research, we have built mathematical framework to analyze some of the approximate inference methods. Using the theoretical results, we have improved the mobile reception of digital satellite broadcasting. Through the interactions within this project, we have also studied the neuron communication channel capacity, the motor planning problem in the brain, and the multi-class classification problem.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
平成18年度	1,500,000	0	1,500,000
平成19年度	2,700,000	0	2,700,000
平成20年度	2,700,000	0	2,700,000
平成21年度	1,600,000	0	1,600,000
年度			
総計	8,500,000	0	8,500,000

研究分野： 総合領域

科研費の分科・細目：情報学・統計科学

キーワード： 数理工学 情報幾何学 信号処理

1. 研究開始当初の背景

(1) 背景：計算機と計測技術の進歩により，現在では大規模なデータに対する高速

な計算手法を用いる必要性が増している。特に推論手法では，ある程度の精度で早く結果を与える近似的解法に注目

が集まりつつあった。

- (2) 動機および準備状況：統計力学の分野では以前より近似解法をあつかっている。それを応用できれば工学的に有用な手法を提案できる。我々はすでに統計力学におけるベーテ近似法に対して、情報幾何学に基づく解析を行っていた。

## 2. 研究の目的

- (1) 統計力学と密接に関係している確率推論の近似的解法の中には応用に結びついていないものがある。一方、情報工学では、近似手法の理論的解析が必ずしも十分ではない。このような確率推論の近似解法について、理論的解析を行なうと同時に、解析結果に基づき情報工学分野での応用を目指す。
- (2) 数理的基盤を発展させる過程で領域内の他の研究班と交流し、新たな研究の創造を目指す。特に応用に関する問題意識の強い情報通信班と理論、応用の相補的立場で協力する。

## 3. 研究の方法

- (1) 情報工学分野での応用として、デジタル衛星放送の移動体における受信を改善する問題を考える。デジタル衛星放送はその歴史的背景から規格そのものが移動体における受信を想定しておらず、自動車などの移動体での受信は簡単ではない。この問題に対して、通信路の推定と推定された通信路に基づく確率推論を時々刻々行うことで受信を改善する。

- (2) 領域内の他の研究班との交流から考えるきっかけを得た問題のひとつとして、多クラス判別機の構成方法について考える。多クラス判別機の構成は一般に難しく、簡単に構成できる2クラス判別機を組み合わせる様々な方法が提案されている。それらの方法を統一的に扱うため、統計モデルとしての一般的な扱い方を考える。

- (3) 領域内の他の研究班との交流から、神経活動の情報伝達を情報理論的な立場から考える。神経細胞同士の情報伝達はスパイクの時間的間隔を通して行われる。このスパイク間隔は正確には制御できず、ノイズが存在することから単位時間に伝達できる情報には限界がある。この量は通信路容量と呼ばれる。神経細胞間の情報伝達における通信路容量を理論的に求める。

- (4) 脳内において、必要な情報をどのように表現しているかは現在様々な面から研究が行われている問題である。特に脳で体の運動を計画する際に、どのように情報表現してい

るのかについて、計算モデルを考える

## 4. 研究成果

(1) まず、デジタル衛星放送の受信について説明する。通信路モデルを考える。実際に車を走行して測定した結果から、次式のような簡単な通信路モデルがよくあてはまることが確認できた。

$$y_t = b_0 x_t + b_1 x_{t-1} + n_t.$$

ここで $y_t$ は時刻  $t$  における観測値、 $x_t$ は送信された信号、 $n_t$ はノイズであり、ここでは正規分布を仮定する。 $b_0$ と $b_1$ は直接波と反射波の強度を示しており、時々刻々と変化する。この通信路モデルをグラフィカルモデルによって表現すると、図1のようになる。

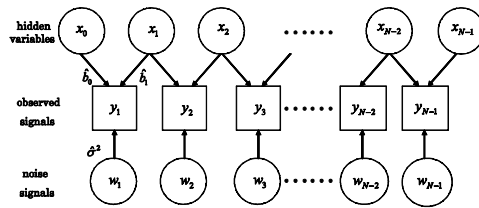


図1 通信路のグラフィカルモデル。

パラメータ $b_0$ と $b_1$ 、ノイズの分散を時々刻々と推定し、その推定されたパラメータを用い、送信された情報を確率伝搬法によって推定する。この計算は簡単であり、十分に実時間での実現が可能である。

現状ではデジタル衛星放送の移動体での受信のためには、高価なアンテナシステムが不可欠であるが、本研究で提案するアルゴリズムを用いれば、安価なシステムによって高品質での受信が可能となる。

表1 提案手法による受信状況の改善。

Route	Average Received CNR [dB]	without Channel Estimation Hard Decision	with Channel Estimation				
			Marginal Inference		Joint Inference		
			Higher Order	EM Algorithm	Higher Order	EM Algorithm	
High Way	H1	9.91	0.00058	<1E-04	<1E-04	<1E-04	<1E-04
	H2	9.47	0.0048	<1E-04	<1E-04	<1E-04	<1E-04
	H3	9.12	0.0063	0.00084	0.00084	<1E-04	<1E-04
	H4	8.88	0.0045	0.0016	0.00040	<1E-04	<1E-04
	H5	8.30	0.028	0.0078	0.0029	<1E-04	<1E-04
Urban	U1	9.39	0.0016	<1E-04	<1E-04	<1E-04	<1E-04
	U2	9.19	0.0058	0.0024	0.00061	<1E-04	<1E-04
	U3	8.40	0.027	0.0071	0.0036	<1E-04	<1E-04
City	C1	9.74	0.0023	0.00045	0.00011	<1E-04	<1E-04
	C2	9.38	0.0021	0.00030	<1E-04	<1E-04	<1E-04
	C3	9.01	0.038	0.0051	0.0051	<1E-04	<1E-04
	C4	8.10	0.016	0.0091	0.0064	<1E-04	<1E-04

表1は提案手法による受信状況の改善を示すものである。パラメータ推定に関して、高次統計量(Higher Order)を用いるものとEM(Expectation-Maximization)アルゴリズムを用いるものを提案した。自動車を高速道路(Highway)、郊外(Urban)、市街地(City)を運転して測定したデータを用いて実験した。Joint Inferenceとして示したものが送信路

を復号した結果である。どの地点であってもエラーが非常に小さく ( $1.0 \times 10^{-4}$ 以下), デジタル衛星放送の受信結果が向上していることがわかる。アルゴリズムは実時間で実現可能であることから, 提案手法が実用上有効であることが示された。

(2) 多クラス判別機の構成方法に関して説明する。2クラス判別機を組み合わせる方法としては ECOC (Error-Correcting Output Coding) に基づく方法と BT (Bradley-Terry) モデルを用いる方法が提案されている。これらは全く別の手法として提案されていることが多いが, その手続きを細かく観察すると, 2クラス判別機を組み合わせる手続きはある非線形凸関数の最適化と同値である。

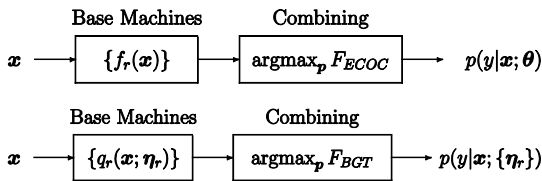


図 2 ECOC 法と BT 法の手続き。

図 2 にあるように, 2つの手法の差は最適化に用いる非線形凸関数  $F_{ECOC}$  と  $F_{BT}$  の差だけである。このことは2つの手法の関係を明らかにしただけではなく, 今後他の手法を提案する際に新たな視点を与えるものである。

(3) 神経細胞間での情報伝達は脳における情報処理の最も基本となる要素である。

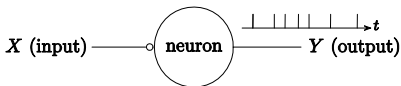


図 3 神経細胞とスパイク列。

神経細胞は時間的に間隔の空いたスパイクと呼ばれる特徴的な信号を送り, 情報を伝達していると考えられるが, 情報を担っているのはそのスパイクの形ではなく, 時間間隔であると考えられている (図 3)。時間間隔は正確に制御できず, ある程度の確率の変動がふくまれているため, 神経細胞間での情報伝達には誤差が含まれる。このような情報伝達路において単位時間にどの程度の情報が伝達できるかという限界を定義したのがシャノンの定義による通信路容量である。

神経細胞同士の情報伝達路においてどの程度の通信路容量をもつかについてはいくつかの結果が知られているが, 用いられている仮定が適切でない, あるいは数学的な証明が十分でないため, さらに議論する必要があった。本研究では計算論神経科学において広く用いられている時間符号化, 頻度符号化の双方を考察し, 通信路容量を与える入力分布

の形が離散分布となることを証明し, 通信路容量を数値的に求めた (図 4, 5)。

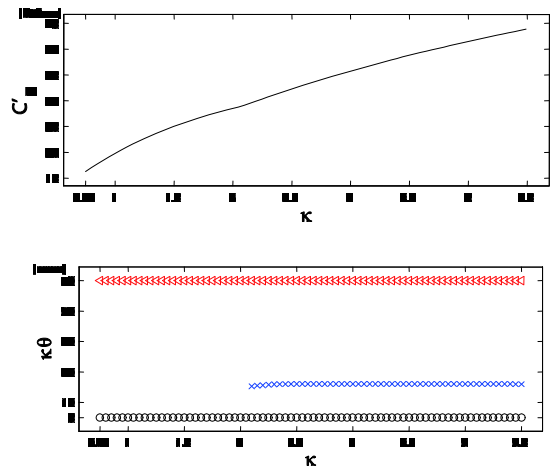


図 4 時間符号化の通信路容量と最適な入力分布。

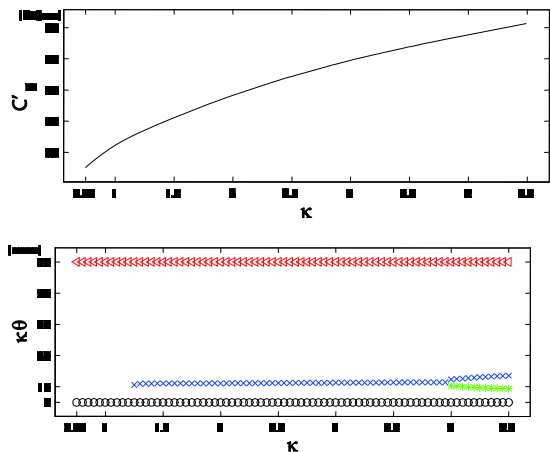


図 5 頻度符号化の通信路容量と最適な入力分布。

(4) 脳における情報処理では, 脳内で情報がどのように表現されているかを考えることが重要である。この研究では, 腕を動かすための信号を脳から筋へ送る際, どのような計算が行われているのかという運動計画の問題を考え, その情報表現を考察した。

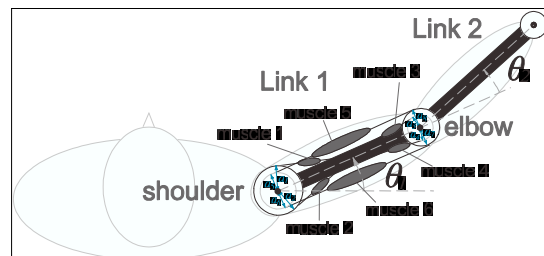


図 6 二関節腕モデル。

図 6 に示した腕モデルを用い, 腕の到達運動を考える。手のある一点から別の一点へ動く軌道を課題としてあたえると, 課題を達成

する運動には特徴がある。まず、目的を達成する軌道は無限に存在するが、人間の手の描く軌道は最短直線の外にふくらみ、速度の時間変化はベル型になる。また、腕の運動に関わる筋の数は冗長であり、この冗長性をどのように解決するかも重要な問題である。到達運動における軌道を説明するモデルは 80 年代より盛んに研究されている。本研究は、それらの研究で十分議論されていなかった脳内での運動指令表現を扱うものである。

ここでは、脳は任意の運動指令を生成できるのではなく、ある基底の線形和の形でしか指令が生成できないとした。この線形和の係数が運動の表現である。運動表現のうち、0 が多い疎な表現を最適な表現だとすれば、筋の冗長性も最適な軌道を求めることも簡単にできる。

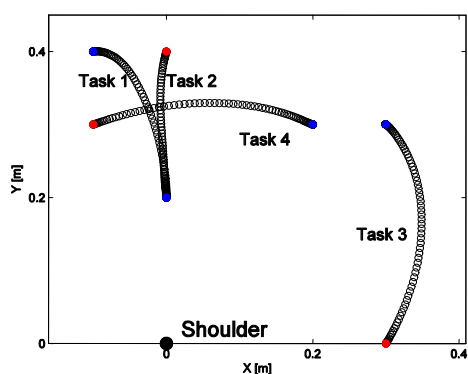


図 7 モデルによって計算された最適軌道.

図 7 はこのような計算モデルによって得られた手の軌道である。赤い点から青い点への到達運動を計算すると、人間の運動の特徴を再現できていることがわかる。速度の時間変化はベル型であり、筋の冗長性も取り除くことができている。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 15 件)

- ① 池田 思朗, Jonathan H. Manton, Capacity of a single spiking neuron, Neural Computation, 21, 1714-1748, 2009, 査読有
- ② 浜田正稔, 池田 思朗, Channel estimation and code word inference for mobile digital satellite broadcasting reception, IEICE Transaction on Communications, E91-B, 3886-3898, 2008, 査読有
- ③ 浜田正稔, 池田 思朗 「通信路推定と誤り訂正による衛星デジタル放送移動受信の改善」 統計数理 第 56 巻 2 号,

pp.215-224, 2008, 査読有

- ④ 阪口豊, 池田 思朗, Motor planning and sparse motor command representation, Neurocomputing, 70, 1748-1752, 2007, 査読有

[学会発表] (計 14 件)

- ① 池田 思朗, 阪口豊, Motor Planning as an Optimization of Command Representation, 48th IEEE Conference on Decision and Control, 2009 年 12 月 23 日, 上海 (中国)
- ② 池田 思朗, Jonathan H. Manton, 神経細胞通信路の通信路容量, 日本神経回路学会第 19 回全国大会, 2009 年 9 月 24 日, 東北大学片平さくらホール
- ③ 池田 思朗, Mean field approximation methods and information geometry, Physics of Algorithms '09, 2009 年 8 月 31 日, サンタフェ (米国)
- ④ 池田 思朗, Jonathan H. Manton, Spiking Neuron Channel 2009 IEEE International Symposium on Information Theory, 2009 年 7 月 1 日, ソウル (韓国)
- ⑤ 浜田正稔, 池田 思朗, Improving mobile reception of digital satellite broadcasting, 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC' 07), 2007 年 9 月 6 日, アテネ (ギリシャ)

[図書] (計 1 件)

- ① 池田 思朗, サイエンス社, 田中和之編著 SGC ライブラリ 50 確率的情報処理と統計力学, 2007 年, pp. 52-58

[産業財産権]

- 出願状況 (計 0 件)
- 取得状況 (計 0 件)

[その他]

研究成果に関する情報は以下のホームページに載せている

<http://www.ism.ac.jp/~shiro/>

## 6. 研究組織

(1) 研究代表者

池田 思朗 (IKEDA SHIRO)

統計数理研究所・数理・推論研究系・准教授

研究者番号 : 30336101

(2) 研究分担者 無

(3) 連携研究者 無