様　式　C－19、F－19－1、Z－19　（共通）

科学研究費助成事業　　研究成果報告書

令和　4 年　 6 月 15 日現在

機関番号：８２６７０
研究種目：基盤研究(C)（一般）
研究期間：2019～2021
課題番号：１９Ｋ０４４５０

研究課題名（和文）条件付き確率密度関数のオンライン学習法とその設計法の確立

研究課題名（英文）An Establishment of On-line Learning Method of Conditional Probability Density Function and Its Design Method

研究代表者
金田　泰昌（Kaneda, Yasuaki）

地方独立行政法人東京都立産業技術研究センター・開発本部情報システム技術部通信技術グループ・上席研究員

研究者番号：２０４６３０１０

交付決定額（研究期間全体）：（直接経費）　　2,200,000円

研究成果の概要（和文）：近年、システムを条件付き確率密度関数としてモデリングする方法が注目されている。しかしながら、既存の方法は多出力系や複雑な分布形状への対応、オンライン学習、データ数削減が個別に検討されおり、全てを考慮した方法は明らかにされていない。そこで、本研究では、多出力系の確率的モデルに対して分布形状の仮定を必要としないオンライン学習法の確立を目的とする。提案法では、システムの確率モデルがカーネルの重み付き平均として与えられると仮定し、ベイズ推論を用いて未知パラメータを求めることで条件付き確率密度関数を導出する。提案法の有効性を数値シミュレーションにより検証する。

研究成果の学術的意義や社会的意義
制御システムを条件付き確率密度関数でモデル化する際、これまではガウス過程回帰やt過程回帰など、パラメトリック分布を仮定した方法がしばしば利用されていた。しかしながら、提案法により、パラメトリック分布を仮定することなく、任意の分布形状の条件付き確率密度関数を求めることができる。このように、条件付き確率密度関数によるモデル化において、統一的な方法を提示した点に学術的な意義がある。また、オンライン学習への拡張が容易であり、学術的な発展性を含んだ成果となっている。

研究成果の概要（英文）：Recently, probabilistic modeling methods, in which systems are represented by conditional probability density functions, are widely noticed. In fields of probabilistic modeling, previous studies proposed modeling methods for multi output systems or non-Gaussian distributions, online learning methods, and methods with a little data. However, these methods are studied individually. In this research, we develop an online probabilistic modeling method for multi output systems without a prior information of probability distributions. A proposed method assumes that a probabilistic model can be written as a weighted average of kernel function. We estimate unknown parameters and derive a conditional probability density function of the model by Bayesian inference. Numerical simulations demonstrated effectiveness of the proposed method.

研究分野：制御工学

キーワード：条件付き密度推定　ノンパラメトリックモデル　ベイズ推定

１．研究開始当初の背景

近年、システムを条件付き確率密度関数としてモデリングする方法（確率的モデリング）が注目されている。しかしながら、既存の確率的モデリングは多出力系や複雑な分布形状への対応、オンライン学習、データ数削減が個別に検討されおり、全てを考慮した方法、すなわち多出力系かつ複雑な分布形状に対応した確率的モデルのオンライン学習法は明らかにされていない。また、既存の確率的モデルの設計法は交差検証による試行錯誤的な方法である。

２．研究の目的

本研究では、多出力系の確率的モデルに対して分布形状の仮定を必要としないオンライン学習法の確立、およびその設計法の確立を目的とする。

３．研究の方法

まずはオフラインにおいて分布形状の仮定を必要としない確率的モデルの学習アルゴリズムを導出する。また、その設計法も合わせて検討する。その後、オフライン学習アルゴリズムを逐次学習アルゴリズムに拡張し、またデータ数削減法を検討する。学習アルゴリズム導出の際の着眼点を(1)に示す。導出したアルゴリズムの有効性は数値シミュレーションで評価する。具体的な評価方法を(2)に示す。

(1) 任意の同時分布をモデル化する方法としてカーネル密度推定（KDE）がよく知られている。例えば、ガウスカーネルを用いたKDEを考えた場合、同時分布$p(x,y)$は次式で推定できる。

$$p(x,y) = \Sigma_{i=1}^{N}\theta_i\mathcal{N}(z|z_i,\Sigma_i),$$

$$\text{where}\quad \Sigma_{i=1}^{N}\theta_i = 1\ .$$

ここで、$\mathcal{N}(z|z_i,\Sigma_i\ )$は平均$z_i$、共分散行列$\Sigma_i$のガウス分布であり、また$z = [x^{T}\quad y^{T}]^{T}$である。なお、通常用いられるKDEは$\theta_i = 1/N$であるが、ここでは一般化した形で定式化する。いま、Σ_iを$\Sigma_i = \begin{bmatrix}\Sigma_{ss}^{i} & \Sigma_{st}^{i}\\ \Sigma_{ts}^{i} & \Sigma_{tt}^{i}\end{bmatrix}$と分割する。このとき、条件付き確率密度関数は次式で与えられる。

$$p(y|x) = \Sigma_{i=1}^{N}\gamma_i\mathcal{N}(t|\alpha_i(x),\Pi_i)\ ,$$

$$\text{where}\quad \Pi_i = \Sigma_{tt}^{i} - \Sigma_{ts}^{i}(\Sigma_{ss}^{i})^{-1}\Sigma_{st}^{i}\ ,\alpha_i(x) = y_i + \Sigma_{yx}^{i}(\Sigma_{xx}^{i})^{-1}(x-x_i),\gamma_i = \frac{\theta_i\mathcal{N}(x|x_i,\Sigma_{xx}^{i})}{\Sigma_{j=1}\theta_j\mathcal{N}\left(x|x_j,\Sigma_{xx}^{j}\right)}.$$

上記で示したとおり、KDEを基礎とした条件付き確率密度関数の導出は、条件付き確率密度関数と比例関係にある同時分布がカーネルの重み付き平均として与えられると仮定するところから議論がはじまる。本研究では、KDEを基礎とした条件付確率密度関数の導出方法を参考に、観測モデル$y = f(x) + v$の確率密度関数がカーネルの重み付き平均として与えられると仮定する。そして、ベイズ推論を用いて未知パラメータを求めることで条件付き確率密度関数を導出する。

(2) 提案法の有効性は数値シミュレーションを用いて評価する。数値シミュレーションでは一例として以下の観測モデルで表されるシステムを考える。

$$y = 5\sin x + v\ .$$

ここで、vは観測ノイズで、以下の三通りのケースを考える。一つめのケースでは、観測ノイズが$\mathcal{N}(v|0,0.2^2)$のガウス分布に従うものとする。二つめのケースでは、位置母数0、尺度母数0.05のコーシー分布に従うものとする。三つめのケースでは、$0.4\mathcal{N}(v|0.5,0.2^2) + 0.6\mathcal{N}(v|-0.5,2\times 0.2^2)$の混合分布に従うものとする。これにより、観測ノイズの確率分布がガウス分布の場合だけでなく、非ガウス分布で表される場合でも提案法が有効であることを確認する。また、確率分布の構造を仮定した方法として、ガウス過程回帰（GP）を用いて観測モデルの確率密度関数を推定した場合と比較し、提案法の有効性を確認する。

４．研究成果

主な成果は以下の(1)および(2)のとおりである。

(1) まず、観測モデルの確率密度関数として、関数値fが与えられたときの観測値yの確率密度関数が次式で表されると仮定する。

$$p(y|f) = \Sigma_{i=1}^{N}\theta_i\mathcal{N}(y|f+v_i,\lambda^{-1})\ . \qquad (1)$$

ここで、f,v_i,λ,θ_iは未知パラメータである。そこで、各パラメータを確率変数とみなし、ベイズ推論によりパラメータの事後分布を計算する。

式(1)はθ_iの確率で要素$\mathcal{N}(y|f+v_i,\lambda^{-1})$が選択されるモデルと等価である。そこで、カテゴリ分布$\mathrm{Cat}(s|\theta)$を用いることで次式のように積の形に書き直すことができる。

$$p(y|f,s,v,\lambda,\theta) = \Pi_{i=1}^{N}\mathcal{N}(y|f+v_i,\lambda^{-1})^{s_i}\ ,\qquad p(s|\theta) = \mathrm{Cat}(s|\theta). \tag{2}$$

ここで、s_iはカテゴリ分布に従う確率変数で、新たに導入された未知変数である。いま、式(2)に対して以下のように事前分布を設定する。

$$p(f) = \mathcal{N}(f|0,K),$$
$$p(v,\lambda) = \mathrm{Gam}(\lambda|a,b)\Pi_{i=1}^{N_k}\mathcal{N}(v\_i|\mu_i,(\beta_i\lambda)^{-1}),$$
$$p(\theta) = \mathrm{Dir}(\theta|\alpha)\ .$$

ここで、$\mathrm{Gam}(\lambda|a,b\ )$は$a,b$をパラメータとするガンマ分布の確率密度関数であり、また$\mathrm{Dir}(\theta|\alpha\ )$は$\alpha$をパラメータとするディリクレ分布の確率密度関数である。このとき、パラメータの事後分布は変分ベイズ法を用いて次式で近似することができる。

$$p(f,s,v,\lambda,\theta) \approx q(f)q(s)q(v,\lambda)q(\theta) \tag{3}$$
$$q(f) = \mathcal{N}(f|\hat{f},\hat{K}),$$
$$q(s) = \Pi_{n=1}^{N}\mathrm{Cat}(s_n|\hat{\eta}_n),$$
$$q(v,\lambda) = \mathrm{Gam}(\lambda|\hat{a},\hat{b})\Pi_{i=1}^{N_k}\mathcal{N}\left(v_i|\ \hat{v}_i,(\hat{\beta}_i\lambda)^{-1}\right),$$
$$q(\theta) = \mathrm{Dir}(\theta|\hat{\alpha}),$$

ここで、＾は変分ベイズ法により計算されたパラメータであり、詳細は割愛するが事前分布で設定したパラメータから計算される。なお、変分ベイズ法を逐次型のマルコフ連鎖モンテカルロ法に置き換えることでオンライン学習が原理的に可能となる。また、同時分布をガウスカーネルの重み付き平均としてモデル化し、事前分布として馬蹄分布やラプラス分布をはじめとするスパースモデルを導入することで係数削減が期待される。

最後に、式(1)および(3)から予測分布として条件付き確率密度関数を導出する。新しい入力x_*に対する関数値および出力値をそれぞれf_*および y_*とする。このとき、式(1)および(3)を用いてパラメータを積分消去することで次式が得られる。

$$p(y_*|f_*) = \Sigma_{i=1}^{N}\hat{\theta}_i\mathrm{St}\left(y_*|f_*+\hat{v}_i,\frac{\hat{\beta}_i\hat{\alpha}}{(1+\hat{\beta}_i)\hat{b}},2\hat{a}\right)\ , \tag{4}$$

ここで、$\mathrm{St}(y|\mu,\lambda,\nu)$はスチューデントの t 分布である。式(4)の$f_*$を積分消去することは困難である。一方、$f_*$は次式のとおり解析的に計算することができる。

$$f_* = k_*^T K^{-1}\hat{f}\ \ . \tag{5}$$

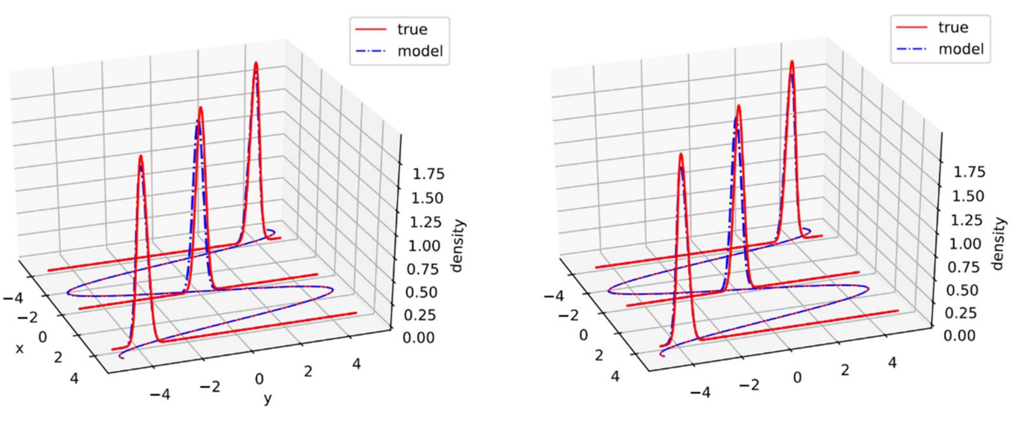
本研究では簡単のため、式(4)に式(5)を代入したものを確率的モデルとして採用する。つまり、

$$p(y_*) = \Sigma_{i=1}^{N}\hat{\theta}_i\mathrm{St}\left(y_*|k_*^T K^{-1}\hat{f}+\hat{v}_i,\frac{\hat{\beta}_i\hat{\alpha}}{(1+\hat{\beta}_i)\hat{b}},2\hat{a}\right)\ . \tag{6}$$

(2) 3．(2)の手順で、4．(1)の提案法の有効性を検証した結果を示す。図 1-3 は、それぞれ観測ノイズがガウス分布、コーシー分布、混合分布に従うものとした場合のモデル化結果を示す。水平面に描かれた曲線が観測モデルの関数値を示しており、縦軸方向の曲線は各関数値が与えられたときの確率密度関数を表す。実線が真値を、点線が推定結果を表す。

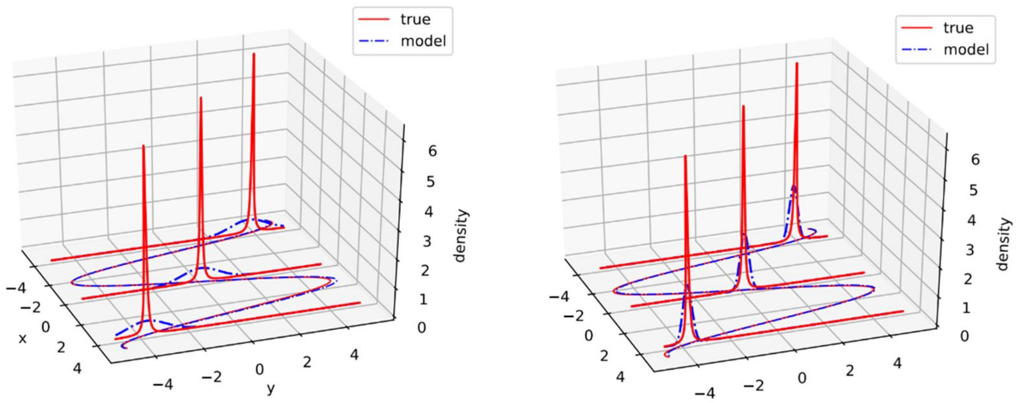
図 1 より、観測ノイズがガウス分布に従うものとした場合は、両者ともに確率密度関数が推定できていることがわかる。一方、図 2 では、提案法により、裾の重さに影響されることなく周度の大きな分布がモデル化できている。また、図 3 では、提案法により多峰性の分布が捉えられていることがわかる。

表 1 に、推定した各分布と真の分布との推定誤差として、分布間のピアソン距離を評価した結果を示す。各値は 10 回計算した際の平均値と標準偏差を表す。この結果からも、提案法は非ガウス分布も含め、広いクラスの分布をモデル化できていることがわかる。



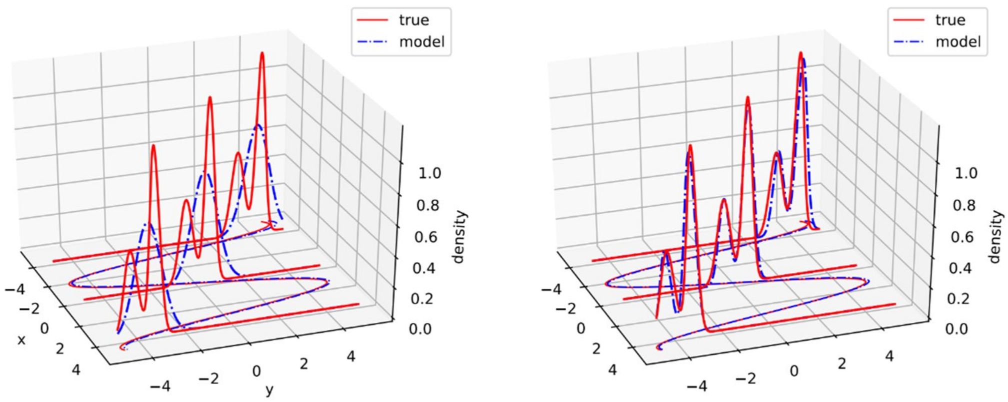
(a)ガウス過程回帰による結果　　　　　　　(b)提案法による結果

図 1：観測ノイズがガウス分布に従う場合の $p(y_k|x_k)$ の推定結果



(a)ガウス過程回帰による結果　　　(b)提案法による結果

図 2：観測ノイズがコーシー分布に従う場合の $p(y_k|x_k)$ の推定結果



(a)ガウス過程回帰による結果　　　(b)提案法による結果

図 3：観測ノイズが混合分布に従う場合の $p(y_k|x_k)$ の推定結果

表 1：ピアソン距離による各分布の推定誤差評価（平均および標準偏差）

	ガウス分布の場合	コーシー分布の場合	混合分布の場合
ガウス過程回帰	0.069(±0.064)	8.310(±5.303)	2.828(±3.090)
提案法	0.084(±0.073)	1.931(±0.701)	0.329(±0.573)

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　1件／うち国際共著　0件／うちオープンアクセス　1件）

１．著者名 Kaneda Yasuaki、Suzuki Satoshi、Irizuki Yasuharu	４．巻 33
２．論文標題 Probabilistic Modeling for Nonlinear and Non-Gaussian Systems Using Kernel Density Estimation and Gaussian Process	５．発行年 2020年
３．雑誌名 Transactions of the Institute of Systems, Control and Information Engineers	６．最初と最後の頁 303～313
掲載論文のDOI（デジタルオブジェクト識別子） 10.5687/iscie.33.303	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計3件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 金田泰昌，入月康晴
２．発表標題 データに基づくシステムの直接設計と粒子フィルタへの応用
３．学会等名 第63回システム制御情報学会研究発表講演会
４．発表年 2019年

１．発表者名 金田泰昌，鈴木聡，入月康晴
２．発表標題 分布形状を仮定しない確率的モデルの学習法
３．学会等名 第62回自動制御連合講演会
４．発表年 2019年

１．発表者名 金田泰昌，鈴木聡，入月康晴
２．発表標題 密度推定を用いた非線形非ガウスシステムのモデル化とベイズ学習
３．学会等名 第64回システム制御情報学会研究発表講演会
４．発表年 2022年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

－

6．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

7．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

8．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関