様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

令和　５ 年　６ 月 １９ 日現在

機関番号：１４４０１
研究種目：基盤研究(C)（一般）
研究期間：2020～2022
課題番号：２０Ｋ１１９８９

研究課題名（和文）確率微分方程式で記述される制御系の有限状態マルコフ鎖近似法の開発とその応用

研究課題名（英文）Development and application of the finite state Markov chain approximation method for control systems described by stochastic differential equations

研究代表者
鈴木　康之（Suzuki, Yasuyuki）

大阪大学・大学院基礎工学研究科・講師

研究者番号：３０６３１８７４

交付決定額（研究期間全体）：（直接経費）　　3,400,000円

研究成果の概要（和文）：本研究では，確率微分方程式で記述される制御システムの初期値・境界値問題を精度よく近似する方法を開発することおよびその応用を目的として，Fokker-Planck方程式を有限要素法に基づいて有限状態マルコフ鎖モデルへ近似する枠組みを構築し，さらに得られた有限状態マルコフ鎖を強化学習やデータ同化問題に適用する基盤整備を進めた．具体例として，確率流が状態依存的に急峻に変化する非線形力学系で記述されるヒト静止立位姿勢間欠制御モデルの有限状態マルコフ鎖近似を行い，本研究で構築した手法が，システムの初期値・境界値問題の数値的近似解を精度良く求められることを明らかにした．

研究成果の学術的意義や社会的意義
Fokker-Planck方程式の動態を近似する数値計算手法は，必ずしも十分に整備されていなかった．本研究でFokker-Planck方程式を有限要素解析に基づき，有限状態マルコフ鎖モデルによって精度よく近似する枠組みの構築したことにより，確率微分方程式で記述される様々なシステムのより詳細な解析への道筋が開けたことになる．さらに本研究では，有限状態マルコフ鎖を強化学習やデータ同化問題に適用する基盤整備を進めた．これにより，確率的に変動する生命現象を様々な手法により取り扱う枠組みの整備が進み，研究の更なる飛躍が期待される．

研究成果の概要（英文）：In this study, we constructed a framework for approximating the Fokker-Planck equation by the finite element method to a finite state Markov chain model, and further developed a foundation for applying the obtained finite state Markov chain to reinforcement learning and data assimilation problems. As a concrete example, we performed a finite state Markov chain approximation of the intermittent control model of human standing posture, which is described by a nonlinear mechanical system in which the probability flow changes steeply depending on the state. We revealed that the method we developed in this study can accurately obtain numerical approximate solutions for initial value and boundary value problems of the system.

研究分野：ソフトコンピューティング

キーワード：確率微分方程式　Fokker-Planck方程式　有限状態マルコフ鎖モデル　強化学習　データ同化

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

連続時間の確率的なシステムの動態は確率微分方程式で記述される．確率微分方程式で記述される制御系の動態解析には，微分方程式を数値積分することで得られる標本道の解析に加え，系の Fokker-Planck 方程式の解として得られる状態点の確率密度関数の時間発展や定常分布を解析することが有用である．以下で記述される，連続時間・連続状態空間の制御システムの確率的なダイナミクス（確率微分方程式）を考える．

$$d\mathbf{X} = \mathbf{f}(\mathbf{X})dt + \mathbf{u}(\mathbf{X})dt + \Sigma d\mathbf{W}$$

ここで $\mathbf{X}$ は実数ベクトル $\mathbf{x}$ をとる確率変数ベクトルであり，$\mathbf{u}(\mathbf{X})$ は制御入力である．$\mathbf{f}(\mathbf{X})$はシステムに制御入力およびノイズが作用しない場合の決定論的ダイナミクス(ベクトル場)を表す．$d\mathbf{W}$ は標準ウィーナー過程の増分，Σ は共分散行列であり，$\Sigma d\mathbf{W}$ でシステムに作用するノイズを意味する．このシステムの確率的なダイナミクスを状態の存在確率分布の時間発展と考え，確率密度関数 $p(\mathbf{x}, t)$の時間発展や定常分布の解析に用いられる Fokker-Planck 方程式（移流拡散方程式）は以下で記述される．

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = -\frac{\partial}{\partial \mathbf{x}}(\mathbf{f}(\mathbf{x})+\mathbf{u}(\mathbf{x}))p(\mathbf{x},t) + \frac{1}{2}\frac{\partial^2}{\partial \mathbf{x}^2}\mathrm{D}(\mathbf{x})p(\mathbf{x},t)$$

ここで右辺第一項はベクトル場による状態の移流を表し，第二項はノイズによる拡散を表す．Fokker-Planck 方程式の数値解法として，空間的な有限差分法がしばしば用いられる．しかしながら，有限差分法を用いて，境界条件付きの偏微分方程式を精度良く解くことは難しい．そこで，Fokker-Planck 方程式に従う確率密度関数 $p(\mathbf{x}, t)$の時間変化を，有限要素法で計算することを考える．有限要素解析では，境界条件を陽に与えて偏微分方程式の数値解を精度よく計算できる．有限要素解析で与える境界条件は，対象として限定された状態空間の境界だけではなく，例えばハイブリッドシステムのように制御入力 $\mathbf{u}(\mathbf{x})$が状態依存的に不連続にスイッチし，移流が激しく変化する境界にも与えられる．Fokker-Planck 方程式の解はマルコフ過程であるため，方程式の有限要素近似を行い離散化した状態点$\tilde{\mathbf{x}}_i$の確率密度$\tilde{p}_i(t) = p(\tilde{\mathbf{x}}_i, t)$の時間変化を表す離散時間・離散状態力学系は，以下の有限状態マルコフ鎖として得られる．

$$\tilde{\mathbf{p}}[t+1] = \mathrm{M}\,\tilde{\mathbf{p}}[t]$$

ここで$\tilde{\mathbf{p}}[t]$は$\tilde{p}_i(t)$を並べたベクトルであり，M はその遷移確率行列である．状態空間の離散化および与えた境界条件が適切であれば，行列演算を逐次的に繰り返すことにより，初期状態からの確率密度関数の時間変化を精度よく計算できる．

以上のように，Fokker-Planck 方程式の有限要素解析を用いれば，確率微分方程式で記述される制御システムのダイナミクスを精度よく計算できると考えられるが，実は，上記の枠組みはこれまで必ずしも十分に整備されていない．その背景には，制御システムの連続的な状態空間を適切に離散化するアルゴリズムや，様々な境界における条件の取り扱いが十分に探求されてこなかったことに原因があると考えられる．一方で，確率微分方程式で記述される制御系のダイナミクスを有限状態マルコフ鎖で近似できると，例えば強化学習やデータ同化など，制御やシステム同定に関する様々な分野への応用が期待できる．

２．研究の目的

本研究では，確率微分方程式で記述される制御システムの初期値・境界値問題を，Fokker-Planck 方程式の有限要素解析に基づき，有限状態マルコフ鎖モデルによって精度よく近似する方法を開発することおよびその応用を目的とする．得られた有限状態マルコフ鎖モデルを，確率微分方程式で記述される制御系の強化学習やデータ同化問題に適用できる可能性を探り，開発した近似モデルの有用性を明らかにする．

３．研究の方法

本研究では，以下に大別する 2 つの課題を実施した：①Fokker-Planck 方程式の有限要素解析に基づく有限状態マルコフ鎖近似の枠組み構築，②有限状態マルコフ鎖近似をデータ同化および強化学習へ応用する基盤の整備．

課題①：Fokker-Planck 方程式の有限要素解析に基づく有限状態マルコフ鎖近似の枠組み構築

Fokker-Planck 方程式を有限要素解析に基づいて離散化し，有限状態マルコフ鎖モデルで近似する方法を開発した．状態空間の離散化方法あるいは境界条件（特にハイブリッド系のスイッチング条件）を適切に扱えなければ，精度よく初期値・境界値問題を近似することができない．本研究では，具体例としてヒト静止立位姿勢の間欠制御モデル（確率的ハイブリッド力学系）を

用いた検証を行った．状態空間の離散化方法および境界条件をさまざまに変えた条件での数値シミュレーションを実施し，それぞれの条件で確率密度関数の時間変化を調べた．この結果と，確率微分方程式のモンテカルロ・シミュレーションを多数回行い，各時刻における状態の確率密度関数を計算した結果を比較した．これにより，本研究で構築した，確率微分方程式で記述される制御システムを Fokker-Planck 方程式の有限要素解析に基づいて精度よく離散状態マルコフ鎖で近似する枠組みの妥当性を検証した．

<u>課題②：有限状態マルコフ鎖近似をデータ同化および強化学習へ応用する基盤の整備</u>

課題①で構築された確率微分方程式で記述される制御システムの離散化および有限状態マルコフ鎖近似をデータ同化および強化学習へ応用する基盤を整備した．

データ同化では，数理モデルと計測データを相補的に融合し，より高度で有用な情報がマイニングされる．あるシステムと，そのシステムのダイナミクスを再現する素質がある「良いモデル」が必要である．本研究では，課題①で用いた間欠制御モデルが良いモデルであることを確認するとともに，有限状態マルコフ鎖をデータ同化へ応用する基盤の整備を目的として，姿勢動揺の計測データと間欠制御モデルのデータ同化を実施した．間欠制御モデルを健常者およびパーキンソン病患者を対象として行った姿勢動揺計測実験データに同化し，ヒト静止立位時の姿勢動揺が間欠制御モデルによりよく再現できることを確認した．

状態空間を離散化して，各姿勢状態依存的に最適な行動を模索する強化学習は，本研究との親和性が高い．本研究で構築した有限状態マルコフ鎖を強化学習へ応用する基盤の整備を目的として，ヒト静止立位の強化学習に関する研究を行った．状態空間を有限要素解析の離散化に倣って離散化し，各離散化状態において“最適”であるとされる行動を Q 学習により学習した．これにより，報酬関数を適切に設定することにより，間欠制御モデルのように，姿勢状態依存的に離散的に最適行動が変化するモデルが得られるかどうかを調べた．

４．研究成果

図 1 に，ヒト静止立位姿勢の間欠制御モデルの Fokker-Planck 方程式を有限要素解析に基づいて離散化し，得られた有限状態マルコフ鎖モデルを用いて確率密度関数の時間変化を数値シミュレーションした結果と，間欠制御モデルのモンテカルロ・シミュレーションを 50,000 回行い，その結果から状態の存在分布の時間変化を計算した結果を示す．本研究では，有限要素解析における境界条件として，境界における確率流束がゼロであるとする（定義域内において確率の合計が一定値である）条件を設定した．対象とする立位姿勢は極めて稀な事象が発生しない限り，平衡点である直立姿勢近傍に姿勢状態を維持するため，この境界条件は妥当である可能性が高い．図 1 に見られるように，各時刻において，状態が存在する確率分布は，本研究で構築した有限状態マルコフ鎖モデルを用いて計算した場合と，モンテカルロ・シミュレーションを行った場合とで定量的に等しい．また，モデルパラメータの値を変えることにより実施した分岐解析においても，両者はほぼ同一の結果を示した．これらの結果は，本研究で開発した，確率微分方程式で記述されるシステムの初期値・境界値問題を Fokker-Planck 方程式の有限要素解析に基づいて近似する枠組みが，妥当であることを意味する．

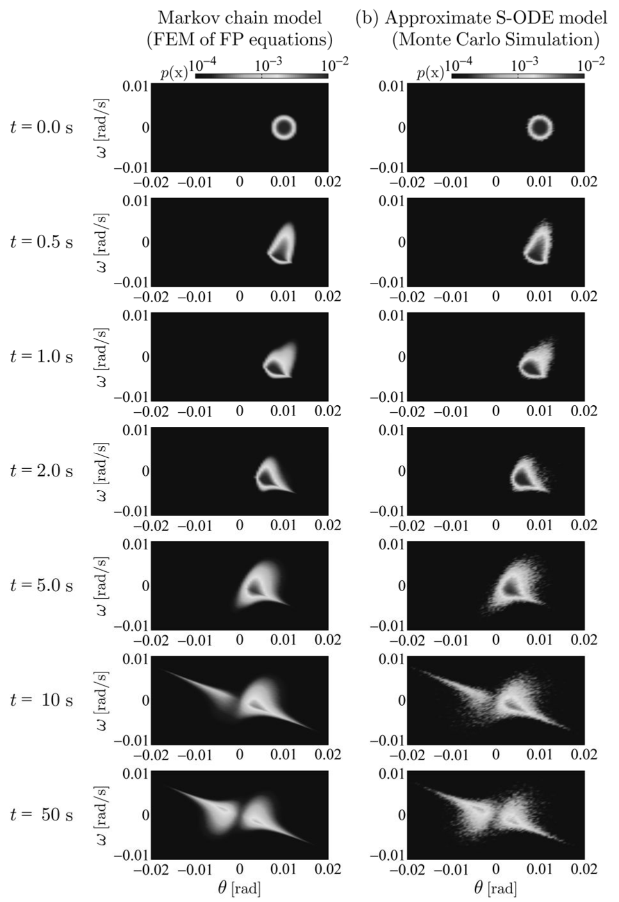

図 1．有限状態マルコフ鎖を用いた確率密度関数の時間変化の数値シミュレーション結果．

図 2 に，ヒト立位姿勢の間欠制御モデルを若年健常者，高齢者，およびパーキンソン病患者から計測した姿勢動揺データに同化し，推定されたモデルパラメータに基づいて実験参加者群を分類した（階層的クラスタリング）結果を示す．図 2 に見られるように，実験参加者は 5 つのグ

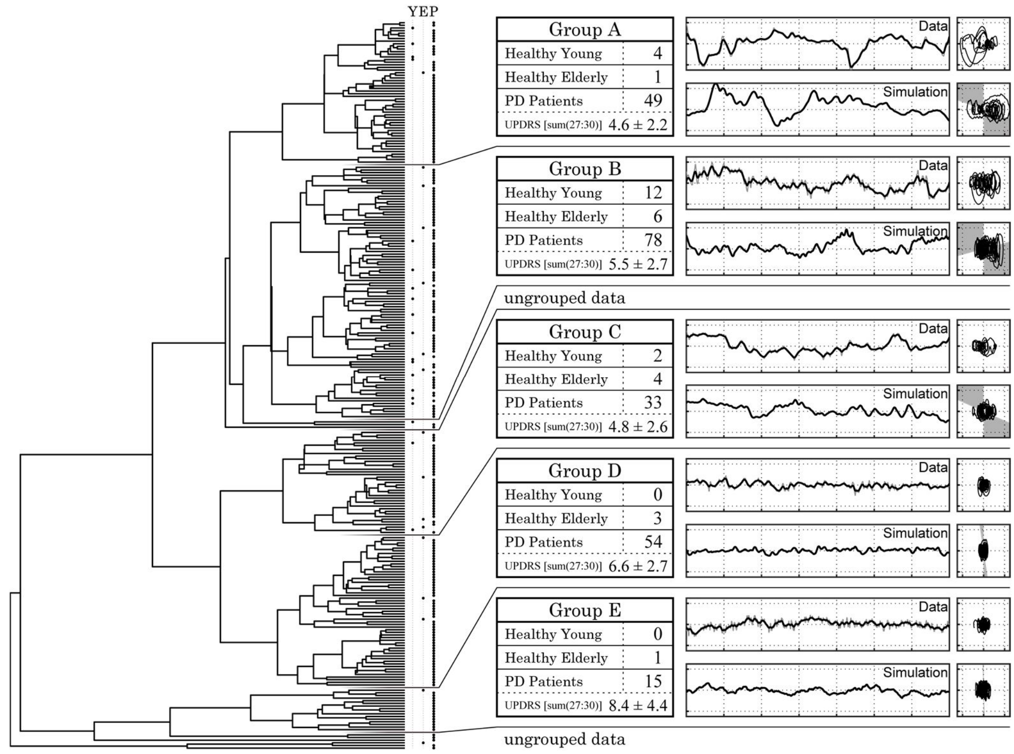

図 2. 間欠制御モデルを姿勢動揺データに同化し，推定されたモデルパラメータに基づいて実験参加者を分類した様子.

ループに大別された．図 2 の右側には，各グループの代表的な実験参加者の姿勢動揺データと，それに基づいて推定されたモデルパラメータを用いて行われたシミュレーションにより得られたモデルの姿勢動揺データを示している．どのグループの実験参加者のデータに関しても，推定されたパラメータを設定することで，間欠制御モデルにより，よく再現できていることがわかる．このことは，ヒト立位姿勢の間欠制御モデルが，ヒトの姿勢動揺を制限する「良いモデル」であることを意味する．

さらに，推定されたモデルパラメータに基づいて各グループを調べると，グループ A-C に分類された実験参加者は立位姿勢を間欠的に制御していたのに対し，グループ D・E に分類された実験参加者は立位姿勢を持続的に制御していた．グループ A-C は主に若年健常者と UPDRS（パーキンソン病重症度の指標）が比較的小さいパーキンソン病患者から構成されているのに対し，グループ D・E は，UPDRS が比較的大きいパーキンソン病患者から構成されていた．これらより，若年健常者は立位姿勢を間欠的な神経制御メカニズムによって維持するのに対し，姿勢の不安定化が見られるパーキンソン病患者は，姿勢制御の間欠性が喪失していることが示された．

強化学習の結果として獲得された方策（姿勢状態において選択するべき行動）の一例を図 3 に示す．定義域が，制御が行われるべき領域（黄色〜赤色）と制御を行わない領域（青色）の領域に大別されていることがよくわかる．このことは，報酬関数の設定によっては，間欠制御と同等な制御戦略が強化学習により獲得されることを意味する．この解析は，有限状態マルコフ鎖近似との親和性が高く，本研究の応用先として，強化学習が大変有効である可能性が示された．

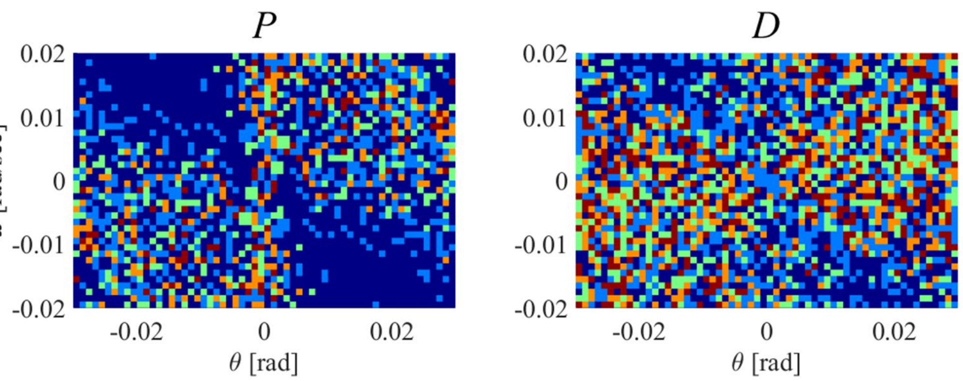

図 3. 強化学習により選択された，各姿勢状態におけるフィードバック制御の比例ゲイン（P）と微分ゲイン（D）の様子.

５．主な発表論文等

〔雑誌論文〕　計2件（うち査読付論文　1件／うち国際共著　1件／うちオープンアクセス　1件）

１．著者名 鈴木康之，十亀敬伍，中村晃大，野村泰伸	４．巻 121
２．論文標題 確率的ハイブリッド力学系の有限要素解析に基づくヒト立位姿勢間欠制御モデルの有限状態マルコフ鎖近似	５．発行年 2021年
３．雑誌名 電子情報通信学会技術研究報告書-IEICE Technical Report-	６．最初と最後の頁 108-113
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 Suzuki Yasuyuki、Nakamura Akihiro、Milosevic Matija、Nomura Kunihiko、Tanahashi Takao、Endo Takuyuki、Sakoda Saburo、Morasso Pietro、Nomura Taishin	４．巻 30
２．論文標題 Postural instability via a loss of intermittent control in elderly and patients with Parkinson’s disease: A model-based and data-driven approach	５．発行年 2020年
３．雑誌名 Chaos: An Interdisciplinary Journal of Nonlinear Science	６．最初と最後の頁 113140～113140
掲載論文のDOI（デジタルオブジェクト識別子） 10.1063/5.0022319	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 該当する

〔学会発表〕　計0件

〔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
研究協力者	中村　晃大 (Nakamura Akihiro)	大阪大学・大学院基礎工学研究科 (14401)	
研究協力者	野村　泰伸 (Nomura Taishin) (50283734)	大阪大学・大学院基礎工学研究科・教授 (14401)	

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