

「確率制約付き変分予測最適制御理論および人間運転行動モデルへの拡張」

瀧 迅（工学研究科），森田 浩（情報科学研究科），高井 重昌（工学研究科）

1. 研究の背景

安全性が重要となる意思決定問題として、医療、経済学、自動運転などの分野が挙げられる。これらの分野では、物理システムの運用において安全要件を考慮することが本質的に必要であり、安全性を無視した場合、人間やその他の対象に対して深刻なリスクを引き起こす可能性がある。したがって、機械学習や強化学習を活用する際には、安全性を考慮しない従来の手法ではなく、学習プロセス自体に安全制約を組み込んだ手法を適用することが不可欠である。このような背景のもと、安全制約付きデータ駆動型変分予測最適制御問題（あるいは安全強化学習）が近年注目を集めており、運用段階での方策の安全性を保証するための有望なアプローチと考えられている。

安全制約付き変分予測最適制御問題は、通常、制約付きデータ駆動型変分予測最適制御問題として定式化され、方策は安全制約を満たすように最適化される。このとき、安全制約の定義には複数の方法があり、適用領域に応じて適切な表現を選択する必要がある。例えば、安全制約の表現には、期待累積安全制約 (Expected Cumulative Safety Constraint)、即時的な厳密制約 (Instantaneous Hard Constraint)、ほぼ確実な安全性 (Almost Surely Safe Constraint)、および結合確率制約 (Joint Chance Constraint) などがある。これらのうち、特に結合確率制約は、ある一定の確率で安全要件を満たすことを保証するものであり、確率的な安全性が求められる環境に適している。

実世界の応用として、ドローンの軌道計画や惑星探査などが挙げられる。これらの

タスクでは、ミッションが有限の時間内で完了する必要があり、その間に安全要件が高い確率で満たされることが求められる。例えば、ドローンが障害物を回避しながら目的地へ向かう際に、衝突リスクを最小限に抑えつつ、確率的に安全な軌道を選択する必要がある。同様に、惑星探査ミッションにおいても、探査機が障害物を避けながら効率的に移動し、ミッションを安全に遂行することが重要となる。このような状況では、決定論的な制約よりも柔軟な表現が求められ、結合確率制約が適用されることで、実用的な方策を導出できる可能性が高まる。

このように、安全性が不可欠な領域では、データ駆動型変分予測最適制御において単に報酬最大化を追求するだけでなく、適切な安全制約を導入することが重要であり、その中でも結合確率制約は、現実世界の応用において特に有用なアプローチとなる。

2. 研究内容と貢献

本研究では、確率制約付きデータ駆動型変分予測最適制御問題 に対して、新たな方策として二値の確率の方策を提案する。本手法では、エージェントが次の行動を決定する際、2つの行動候補の間で確率的な選択を行う。具体的には、各状態ごとに設定された2つの行動候補のうち、確率的にどちらかを選択する仕組みであり、この確率は状態に依存して変化する。この確率的な選択は、理想的なコインではなく、歪みを持つ可能性のあるコインを投げる操作になぞらえることができる。すなわち、各状態において、適切に設定された確率に基づき、2

つの候補のうち一方の行動が選択される。このような確率的な方策は、従来の決定論的方策とは異なり、柔軟な意思決定を可能にする。

本研究では、まず、CCMDP に対するベルマン方程式を導出し、さらに、最適解集合内に二値の確率的方策が存在することを理論的に証明する。これは、確率制約付き問題において、単純な決定論的方策だけでなく、確率的な方策の導入が最適解の範囲を広げることが示す重要な結果である。しかしながら、CCMDP において直接的に結合確率制約 (Joint Chance Constraint) を満たす最適方策を求めることは、計算的に非常に困難である。結合確率制約は、複数の状態・時間にまたがる制約であり、解析的に取り扱うことが難しいため、そのままの形で最適化を行うことは現実的ではない。

そこで、本研究では、結合確率制約を期待累積安全制約 (ECSC) に近似する手法を提案する。この近似によって、確率制約の解析が容易になり、より実用的な最適化手法が適用可能となることを示す。さらに、我々の理論的解析により、ECSC を課した制約付き MDP においても、最適解集合内に二値の確率的方策が存在することを証明する。すなわち、本研究で提案する二値の確率的方策は、結合確率制約を直接扱う場合だけでなく、その近似問題である期待累積安全制約を用いた問題設定においても、有効な解の一部として存在することが保証される。

最後に、Safety Gym ベンチマーク において、我々の提案する二値の確率的方策が、既存の安全制約付きデータ駆動型変分予測最適制御アルゴリズムの性能を向上させることを示す。本手法は、同じ安全制約の条件下において、従来手法と比較してより優れた性能を達成することを確認した。この結果は、二値の確率的方策が、安全制約付

きデータ駆動型変分予測最適制御において有効なアプローチであることを示唆している。

以上の結果から、本研究で提案する二値の確率的方策は、確率制約付きデータ駆動型変分予測最適制御問題の文脈において、実装の容易さと理論的な正当性を両立する有望なアプローチとなる。特に、確率的な方策を導入することで、制約を満たしながら柔軟な意思決定が可能となり、安全性が求められる応用（例：自動運転、ロボット制御、金融リスク管理など）において有効な手法となることが期待される。

本研究の最大の貢献は、確率制約付きマルコフ決定過程 (CCMDP) において、二値の確率的方策が最適解として機能することを理論的に証明し、その学習手法を確立したことである。これにより、確率制約を考慮したデータ駆動型最適制御に新たなアプローチを提供し、安全性が求められる多くの実応用において実用性の高い意思決定手法を提案することが可能となった。

(a) 確率制約付きデータ駆動型変分予測最適制御問題における新たな方策の提案：従来の確率制約付きデータ駆動型変分予測最適制御問題では、決定論的方策や標準的な確率的方策に依存しており、確率制約を満たしながら最適方策を求めることが困難であった。本研究では、二値の確率的方策を導入することで、制約付き最適化問題に対する柔軟な解法を提供する。この方策では、2つの行動候補の間で確率的な選択を行い、状態依存のフリップ確率に基づいて行動を決定する。これにより、確率制約を満たしつつ、より実用的な方策の設計が可能となる。

(b) CCMDP に対する新たな理論の確立：本研究では、確率制約付き MDP に対するベルマン方程式を導出し、二値の確率的方策が最適解として存在することを厳密に証明

した。これは、従来の制約付きデータ駆動型変分予測最適制御の枠組みでは十分に明らかにされていなかった重要な性質であり、確率制約を持つデータ駆動型変分予測最適制御問題の解法に新たな理論的基盤を提供するものである。この結果により、確率制約を満たす最適方策を求めるための新たな手法が確立され、確率的安全性を考慮したデータ駆動型変分予測最適制御における理論的な理解が深まる。

(c) 結合確率制約の近似理論の確立: 確率制約付きデータ駆動型変分予測最適制御問題における主要な課題の一つは、結合確率制約 (Joint Chance Constraint) を直接扱うことが計算的に困難である点である。本研究では、結合確率制約を期待累積安全制約 (ECSC) で保守的に近似できることを示し、この近似が最適解の構造を大きく損なわないことを理論的に証明した。これにより、従来は困難であった結合確率制約付き MDP の最適化問題を、より計算可能な問題に変換することが可能となった。

(d) 既存のデータ駆動型変分予測最適制御アルゴリズムとの統合: 本研究では、期待累積安全制約 (ECSC) を用いることで、既存のデータ駆動型変分予測最適制御アルゴリズム (例えば、Constrained Policy Optimization (CPO)) を適用可能であることを示し、それに基づく学習フレームワークを提案した。これにより、従来の制約付きデータ駆動型変分予測最適制御手法と互換性を保ちつつ、新たに提案した二値の確率的方策を効率的に学習することが可能となった。特に、本手法を CPO に適応させることで、最適な二値の確率的方策の近似解を求めることができることを示した。

(e) 実験による有効性の検証: 本研究では、Safety Gym ベンチマークを用いた実験を行い、提案手法が既存のデータ駆動型変分予測最適制御アルゴリズムと比較して優れ

た性能を達成することを示した。特に、同じ安全制約の条件下で、提案手法はより高い報酬を達成しながら、安全性を維持できることを確認した。これは、二値の確率的方策が、実際のデータ駆動型変分予測最適制御タスクにおいて安全性と性能のトレードオフを最適化する上で有効であることを示唆している。

まとめ: 本研究は、確率制約付きデータ駆動型変分予測最適制御問題の分野において、理論的な進展と実用的な手法の両面から重要な貢献を果たすものである。特に、二値の確率的方策の導入とその最適性の証明、結合確率制約の近似理論の確立、既存のアルゴリズムとの統合、そして実験による性能向上の確認は、本手法の有効性を強く示唆する。今後、本手法をさらに発展させ、より複雑な環境や大規模な実応用に適用することが期待される。

3. 人間運転行動モデルへの拡張

人間ドライバーの運転行動を説明可能な形でモデル化するパラメータ化された変分予測最適制御フレームワークを提案する。提案の人間行動モデルには、社会的力ダイナミクスを組み込んだ変分予測最適制御を用いて、車両と他の交通参加者との相互作用を考慮し、現実的な運転行動を再現する。さらに、人間の運転問題を制約付き変分予測最適制御問題として定式化し、最適方策が制約付き変分予測最適制御の最適方策と同値であることを証明する。また、粒子群最適を用いて実データから個々の運転スタイルを学習する手法を開発し、シミュレーションおよび実験により、提案手法の精度と実用性を示す。本研究は、自動運転の安全性検証やシミュレーションテストへの応用が期待される。

引用文献

該当なし

発表論文等

〔雑誌論文〕

- ① Xun Shen, Ye Wang, Kazumune Hashimoto, Yuhu Wu, Sebastien Gros, "Probabilistic reachable sets of stochastic nonlinear systems with contextual uncertainties," Automatica, Accepted as Regular Paper.
- ② Kazumune Hashimoto, Yuga Onoue, Akifumi Wachi, Xun Shen, "Learning-based event-triggered MPC with Gaussian processes under terminal constraints," IEEE Transactions on Cybernetics, Accepted as Regular Paper.
- ③ Nan Yang, Zhendong Xiong, Li Ding, Yikui Liu, Lei Wu, Zhao Liu, Xun Shen, Binxin Zhu, Zhenhua Li, Yuehua Huang, A game-based power system planning approach considering real options and coordination of all types of participants, Energy, vol. 312, no. 133400, 2024. DOI: 10.1016/j.energy.2024.133400.
- ④ Xun Shen, Set-valued regression of wind power curve, IEEE Transactions on Sustainable Energy, Early Access, DOI: 10.1109/TSTE.2024.3458916.
- ⑤ Xun Shen, Hampei Sasahara, Jun-ichi Imura, Makito Oku, Kazuyuki Aihara, Re-stabilizing large-scale network systems using high-dimension low-sample-size data analysis, IEEE Transactions on Emerging Topics in Computational Intelligence, Early Access, DOI: 10.1109/TETCI.2024.3442824.
- ⑥ Xun Shen, Tinghui Ouyang, Kazumune Hashimoto, Yuhu Wu, Sample-based continuous approximate method for

constructing interval neural network, IEEE Transactions on Neural Networks and Learning Systems, Early Access, doi: 10.1109/TNNLS.2024.3409379.

- ⑦ Wataru Hashimoto, Kazumune Hashimoto, Akifumi Wachi, Xun Shen, Masako Kishida, Shigemasa Takai, Bayesian meta-learning on control barrier functions with data from on-board sensors, Advanced Robotics, Early Access, DOI: 10.1080/01691864.2024.2401897.

〔学会発表〕

- ① Xun Shen, Shuo Jiang, Akifumi Wachi, Kazumune Hashimoto, Sebastien Gros, Flipping-based policy for chance-constrained Markov decision processes, Proceedings of the 38th International Conference on Neural Information Processing Systems (NeurIPS 2024), Accepted.
- ② Ye Wang, Xun Shen, Hongyu Qian, Stochastic model predictive control with probabilistic control barrier functions and smooth sample-based approximation, The 2024 63rd IEEE Conference on Decision and Control (CDC), Accepted.
- ③ Xun Shen, Hampei Sasahara, Jun-ichi Imura, Kazuyuki Aihara, "Small-sample-size data-driven disease-detection and re-stabilization for mRNA-protein gene regulatory networks," the 2025 American Control Conference (ACC), DENVER, Colorado, USA, July 7-10th, 2025.

〔その他〕（招待講演）

- ① Xun Shen, Chance constrained optimization for wind power curve fitting with unclean data, The 8th

NUS-ISM-ZIB-IIR-MODAL Workshop:
Next Generation Computing and
Algorithms in the Digital Era, August 12,
2024.