

(続紙 1)

京都大学	博士 (情報学)	氏名	張 靖宇 (Zhang Jingyu)
論文題目	Fast Algorithms for Stochastic Model Predictive Control with Chance Constraints via Policy Optimization (方策最適化による機会制約付き確率モデル予測制御の高速アルゴリズム)		
(論文内容の要旨)			
<p>Stochastic model predictive control (SMPC) is currently receiving increasing attention in the control of stochastic systems because it provides a systematic way to incorporate probabilistic descriptions of uncertainties in a stochastic optimal control problem. The major challenge of SMPC comes from finding an optimal policy at each time instant. The goal of this thesis is to find novel policy optimization methods for SMPC with chance constraints, which can handle different types of problem settings, and obtain good closed-loop control effects and computational efficiency.</p> <p>Chapter 1 gives a comprehensive overview of model predictive control (MPC) and SMPC. It then describes challenges in realizing SMPC with chance constraints.</p> <p>Chapter 2 introduces preliminary knowledge of solving SMPC, which is used throughout this thesis. The first part of this chapter introduces the basic knowledge of numerical optimization, and the second part introduces some important features of SMPC.</p> <p>Chapter 3 presents an efficient control parameterization method for linear SMPC with chance constraints. This parameterization combines the related simplification techniques of affine disturbance feedback to create a simplified affine disturbance feedback parameterization. The number of decision variables is decreased to grow linearly with respect to the horizon length compared with quadratic growth of the original affine disturbance feedback control law, resulting in a preferable trade-off between real-time calculation and control performance. This parameterization is shown to be equivalent to a state feedback control law, and the closed-loop stability of the SMPC problem can also be guaranteed under mild assumptions. The simulation results show that the proposed control parameterization method provides a desirable control performance with low computation cost and achieves the expected result.</p> <p>Chapter 4 presents a numerical algorithm for linear SMPC with chance constraints via solving the Bellman equation. The proposed approach reformulates the SMPC problem in a stochastic programming fashion. A recursive Riccati interior-point method is proposed to solve the ensuing inequality-constrained dynamic programming. The proposed method eliminates active sets in conventional explicit MPC and does not suffer from the curse of dimensionality because it finds the value function and feedback policy only for a given state using the interior-point method. Moreover, the proposed method is proven to converge globally to a stationary solution Q-superlinearly. The numerical experiment reveals that the proposed method achieves a less conservative performance with low computational complexity compared to existing methods.</p> <p>Chapter 5 presents an efficient numerical algorithm for output-feedback nonlinear SMPC with chance constraints. The stochastic optimal control problem is also solved in a stochastic</p>			

dynamic programming fashion, like in Chapter 4, and the output-feedback control is performed with the extended Kalman filter. The information state is summarized as a dynamic Gaussian belief model. Thus, the stochastic Bellman equation is transformed into a deterministic equation using this model. A novel constrained approximate dynamic programming algorithm is proposed to solve the resulting constrained Bellman equation. The proposed algorithm is proven to exhibit a Q-superlinear local convergence rate. The numerical experiment shows that the proposed method achieves good control performance and a reasonable level of constraint violation and is computationally efficient owing to the Riccati-type structure.

Chapter 6 presents a sample-based Bayesian reinforcement learning method for dealing with systems of unknown models. The unknown model is learned by a Gaussian process dynamic model in a model-based Bayesian reinforcement learning framework. The partially observable Markov decision process is reformulated as a belief Markov decision process by the particle filter. Using a sample-based method, the stochastic dynamic programming is transformed into several deterministic constrained dynamic programming problems. Each deterministic problem is solved by the algorithm proposed in Chapter 5. The chance constraint is treated by a sample-removal algorithm. The numerical experiment shows that the proposed method can get good control performance with a reasonable level of constraint violation. Compared with the conventional Bayesian reinforcement learning method, the learning efficiency is significantly improved.

Chapter 7 summarizes this thesis and discusses the directions of future work.

(論文審査の結果の要旨)

確率モデル予測制御 (Stochastic Model Predictive Control: SMPC) は、プロセスシステム、電力システム、ファイナンス、自動運転システムなど、不確実性を含むさまざまな動的システムを安全かつ最適に運用するためのフィードバック制御手法である。SMPCでは、動的システムにおける確率的な不確かさと機会制約 (与えられた不等式が成り立つ確率に対する制約) を考慮し、各時刻で有限時間未来までの評価関数を最小化して制御入力を決定する。その際、時間関数であるフィードフォワード制御入力を最適化するより、状態の関数であるフィードバック制御方を最適化した方が性能は向上することが知られている。しかし、フィードバック制御方の最適化は一般に困難であり、従来の数値計算アルゴリズムには計算量や記憶量、近似精度に改善の余地があった。本論文は、SMPCのいくつかの問題設定に対して高速なアルゴリズムを提案している。具体的には、以下に示す研究成果を得ている。

- (1) 線形 SMPC において最適化するフィードバック制御方の表現として、簡略化アフィン外乱フィードバック表現を提案している。このように表現されたフィードバック制御方の最適化が、あるクラスの状態フィードバック制御方の最適化と等価であり、かつ、凸最適化問題に帰着できることを示している。数値例によって、低い計算コストと高い制御性能が両立できることを示している。
- (2) 線形 SMPC のための再帰的リッカチ内点法を提案している。問題の構造を利用することによって、確率動的計画法を確定的な動的計画法に帰着させ、さらに、現在の状態から出発する軌道に沿った局所的な状態フィードバック方がリッカチ型の方程式と内点法の組み合わせで計算できることを示している。提案手法によって、現在の状態における制御入力を評価区間長に関して線形オーダーの計算量で求めることができる。提案手法の超線形収束を示すとともに、数値例によって、極めて低い計算コストで高い制御性能が実現できることを示している。
- (3) 非線形出力フィードバック SMPC のための近似動的計画法を提案している。軌道に沿った不確かさの伝播も考慮した反復計算によって、未来の外乱を考慮した局所的なフィードバック制御方の最適化を実現している。提案手法の超線形収束を示すとともに、数値例によって、極めて低い計算コストで高い制御性能が実現できることを示している。
- (4) 未知システムに対するベイズ型強化学習法を提案している。未知システムのモデルをガウス過程回帰によって学習し、不確かさを多数のサンプルによって表現することで、確率動的計画法を確定的な動的計画法に帰着させている。その結果、本論文で提案した近似動的計画法が適用可能になっている。数値例によって、非ガウス性外乱を持つ未知非線形システムに対して機会制約を満たす最適なフィードバック方が設計できることを示している。

以上を要するに、本論文は、動的システムにおける確率的な不確かさを陽に考慮して制御性能を向上させるSMPC問題を高速に解く新しいアルゴリズムの体系を開発し、動的システムのより高性能かつ安全な制御の可能性を切り拓いている。よって、本論文は博士 (情報学) の学位論文として価値あるものと認める。また、令和5年2月16日、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。本論文のインターネットでの全文公表についても支障がないことを確認した。