

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	ロボットラーニング (松原 崇充 (教授))		
学籍番号	2111292	提出日	令和 6年 1月 18日
学生氏名	米澤 壮太郎		
論文題目	分布ドリフトを伴う継続学習のためのリザーバサンプリング		
要旨			
<p>近年、ロボットや自動運転に対して、不確実な環境に適用可能なモデル予測制御やモデルベース強化学習による制御が注目されている。これらの手法では、正確なダイナミクスモデルを必要とする。そこで、関数近似性能の高いニューラルネットワークを用いて、観測データに合致するようにモデルを学習する方法が有効とされる。</p> <p>しかし、幅広い環境で活動するロボットを想定すると、あらゆる環境における観測データを事前に用意することは困難である。そこで、逐次的に観測データを獲得しながらモデル学習を継続的に進める必要がある。このとき、ニューラルネットワークで表現したモデルでは、以前学んだ知識を忘れてしまう、破滅的忘却と呼ばれる問題が生じる。これを緩和して学習した知識を定着させるべく、過去の全観測データを一樣確率でメモリに保持し、学習に再利用するリザーバサンプリングが多くの継続学習手法で用いられている。</p> <p>一方で、実世界における環境やロボットは非定常であり、人の動き方、気温、照明、ロボットの摩耗など、様々な不可観測要因によって観測データの生成分布が変化する、データ分布ドリフトとして知られている。新たな生成分布へとモデルを適応すべく、ドリフト発生時に最新のデータから優先的に学習する、ドリフト適応手法が有効である。</p> <p>以上のように、過去のデータからの学習を要請する継続学習と、最新のデータからの学習を要請するドリフト適応は、トレードオフの関係にあり両立が難しい。しかし、どちらも実世界でロボットが学習・活動するためには必要不可欠な能力である。そこで、本研究では、継続学習とドリフト適応の両立を可能にする手法を開発し、継続的かつ適応的なロボットのモデル学習を目指す。</p> <p>継続学習とドリフト適応を両立させるべく、リザーバサンプリングにおける観測データのメモリへの保持確率に着目する。一般的には、過去の全観測データを一樣に保持するように設計されているが、これは同時に最新のデータを優先してドリフト適応することを阻害してしまう。そこで、新たに設計・一般化した保持確率により、最新のデータの優先度、ひいては過去のデータの廃棄速度を調整可能にする。この新たなリザーバサンプリングを適切に用いて、継続学習とドリフト適応を両立可能な塩梅にしてトレードオフを解消する。</p> <p>提案手法の有用性を検証するために、破滅的忘却とデータ分布ドリフトの両方を抱える環境での比較実験を実施する。まず、人工的な回帰問題において、提案手法が指定した保持確率に応じて継続学習能力とドリフト適応能力を増減・両立可能であることを示す。そして、ロボットの動力学シミュレーションにおいて、モデル学習における予測精度を評価する。</p>			