

2025/07/23 分野抄読会

正則化を用いた強化学習的動的治療レジメンモデル

東北大学大学院 医学研究科 公衆衛生学専攻
医学統計学分野 修士一年 矢澤清麻

概要

実際の臨床状況においては、患者の多様な背景や時間経過に伴う病態変化に適応した動的治療レジメンの必要性が高まっている。特にがん治療や精神疾患の薬物療法などは多種多様な治療選択肢が存在し、個人差や時間差を考慮した動的な治療設計が求められる。このような動的な治療設計の効果を求める手法には g-estimation や構造化ネスト平均モデルなどの因果推論の理論などがある。しかしこれらの手法は因果効果を推定できる一方、個人差や時間差を考慮することや、動的に治療設計を探索ができないという問題がある。それに対して強化学習 (Q-Learning) とは、マルコフ決定過程をベースとし、現在の状態からどの行動 (治療) をとれば、長期的な累積報酬 (治療効果) を最大化できるかを探索するのに適した手法である。強化学習を用いれば個人差、時間差を考慮し、動的な治療設計の探索が可能となる。しかし、通常の強化学習は高次元の特徴空間での計算コスト増大、サンプルが少ない場合の推定不安定性、過学習といった課題を抱えている。そこで Penalized Q-Learning では L1 正則化を用いることで過学習といった課題を解決し、それに加えて治療効果のない個人の特定を行っている。

本抄読会では Penalized Q-Learning と正則化について紹介し、今後の展望としてがん性疼痛におけるオピオイドにおける適用を考察した。