THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

Master’s Thesis Presentation

Xinyue Lou
Department of Statistics
The University of Chicago

“Low-Rank Reinforcement Learning for Robust Value Estimation”

November 12, 2025, at 12:00 PM
Jones 111, 5747 S. Ellis Avenue

Abstract

Offline deep reinforcement learning (RL) must learn reliable value functions from a fixed, imperfect dataset, where distribution shift induces extrapolation error and noisy Bellman targets compound through bootstrapping. While conservative methods such as CQL mitigate out-of-distribution (OOD) overestimation by penalizing large Q-values, they frequently become over-pessimistic—shrinking in-distribution values, requiring careful penalty tuning, and limiting expressiveness.

This thesis introduces Structured Value-based Reinforcement Learning (SVRL)and Uncertainty-Aware Low-Rank Q Estimation (UALQE), two complementary approaches that inject a low-rank inductive bias into value learning. At each update, we view the batch-wise next- state/policy action cross-product as a partially observed matrix and apply soft-impute re-construction to exploit shared structure across state–action pairs. We operationalize this reconstruction through T-matrix replacement that denoises TD targets via reconstructed diagonals. UALQE replaces random masking with uncertainty-guided masking (via an ensemble/EDAC-style diversification), concentrating reconstruction on the most error-prone entries.

AnSVDtrackerrevealsthatSVRL/UALQEstabilizethespectrumoflearnedQ-matrices: unlike SAC’s rank-one collapse and CQL’s capped but over-compressed spectra, our methods preserve a compact yet multi-modal structure (moderate $\sigma 1$, elevated approximate rank), with dominant singular vectors exhibiting coherent, non-uniform patterns. Across standard D4RL continuous-control tasks, SVRL attains the most consistent gains, and UALQE reliably outperforms random masking. These results show that structural low-rank reconstruction, paired with uncertainty-

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@uchicago.edu. If you wish to subscribe to our email list, please visit the following website: https://lists.uchicago.edu/web/info/statseminars.

aware selection, provides a practical alternative to uniformly pessimistic objectives, mitigating over/under-estimation while improving stability and return under distribution shift.

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@uchicago.edu. If you wish to subscribe to our email list, please visit the following website: https://lists.uchicago.edu/web/info/statseminars.