



Open Thesis (MA)

Policy Iteration and Value Iteration for Discrete-Time Infinite-Horizon LQR

Description:

One of the main methods to solve optimal control problems is **dynamic programming (DP)**. Various successful algorithms, including **Q-function value iteration (Q-VI)**, **Q-function policy iteration (Q-PI)** and **value-function policy iteration (V-PI)**, have been developed to solve DP problems. Many Reinforcement learning algorithms are based on these algorithms. However, there is not in the literature a unified view on these schemes and a precise characterization of when they work. The main idea of this thesis is to conduct analytical comparisons of the Q-VI, Q-PI and V-PI algorithms in the context of the discrete-time infinite-horizon linear quadratic regulator problem. The student's tasks include conducting a literature review on these algorithms and subsequently delving into the theoretical aspects of these algorithms to provide analytical insights into the conditions of convergence rate and robustness for each algorithm. The final stage aims to uncover the high-level connections between these algorithms, supported by some simulations.

Prerequisites:

- Knowledge in *Optimal Control (Dynamic Programming)*
- Interest in theoretical problems
- Programming skills, *Matlab or Python*

Supervisor:

Bowen Song
Room 2.236

Area:

Optimal Control

Properties:

Type: **MA**

30% literature
50% theory
20% simulation

Beginning:

Any time

Further information on www.ist.uni-stuttgart.de/lehre/bama

Aushang vom December 1, 2023