
Examination

Theoretical examination

1 QUESTIONS

- What is a policy? A stationary policy? An optimal policy? Can we only consider stationary policies to find an optimal one? Justify your answer.
- What is supervised batch-mode learning? How is Fitted-Q-Iteration related to batch-mode learning?
- Describe the Fitted-Q-Iteration (FQI) algorithm. Give examples of supervised learning algorithms that can be used inside FQI. Describe situations where the sequences of Q_N -functions computed with FQI (i) can diverge, (ii) converge and (iii) do not converge but still lead to high quality policies.
- Give two possible practical stopping conditions for FQI and explain their drawbacks.
- Give the three classical results from dynamic programming theory that FQI exploits.
- Can a sequence of Q_N -functions, computed with FQI and tree-based supervised learning, diverge to ∞ ? If not, provide bounds.