Fitted Q-iteration in continuous action-space Markovian Decision Processes

András Antos, Remi Munos, Csaba Szepesvári

Poster ID T6

Fitted Q-iteration:

$$
\begin{aligned}
Q_{k+1} &= \mathrm{Regress}(D_k(Q_k)), \\
D_k(Q_k) &= \left\{ \left[(X_t, A_t), R_t + \gamma \max_{b \in \mathcal{A}} Q_k(X_{t+1}, b) \right]_{1 \le t \le N} \right\}.
\end{aligned}
$$

- Does this algorithm work?
- When?
  - Rigorous analysis
  - Actor-critic variant
  - Error bounds

Continuous action spaces

Continuous state spaces

Single trajectory input