Lower Bounds for RL

Jingfeng Wu

Created: July, 2020
Last updated: August 22, 2020

1 Preliminary

1.1 Yao's principle

Theorem 1 (Yao's principle)**.** *Let X be a finite set of inputs, $\mathcal{A}$ be a finite set of deterministic algorithms. Let R be a random algorithm, i.e., an algorithm that is randomly drawn from set $\mathcal{A}$ according to certain distribution $\mathcal{R}$. Let $\mathcal{D}$ be a distribution over the input set X. Let $\mathrm{cost}(\mathcal{R}, x) = \mathbb{E}_{A\sim\mathcal{R}}[\mathrm{cost}(A, x)]$. Let $\mathrm{cost}(A, \mathcal{D}) = \mathbb{E}_{x\sim\mathcal{D}}[\mathrm{cost}(A, x)]$. Then we have*

$$\min_{\mathcal{R}} \max_{x\in X} \mathrm{cost}(\mathcal{R}, x) = \max_{\mathcal{D}} \min_{A\in\mathcal{A}} \mathrm{cost}(A, \mathcal{D}).$$

Corollary 1.1.

$$\max_{x\in X} \mathrm{cost}(\mathcal{R}, x) \geq \min_{A\in\mathcal{A}} \mathrm{cost}(A, \mathcal{D}).$$

1.2 Fano's inequality

2 Regret lower bound

The lower bound is due to [Auer et al., 2009].

3 PAC lower bound

The lower bound is due to [Azar et al., 2013].

4 PAC lower bound for reward-free exploration

The lower bound is due to [Jin et al., 2020].

References

P. Auer, T. Jaksch, and R. Ortner. Near-optimal regret bounds for reinforcement learning. In *Advances in neural information processing systems*, pages 89–96, 2009.

M. G. Azar, R. Munos, and H. J. Kappen. Minimax pac bounds on the sample complexity of reinforcement learning with a generative model. *Machine learning*, 91(3):325–349, 2013.

C. Jin, A. Krishnamurthy, M. Simchowitz, and T. Yu. Reward-free exploration for reinforcement learning. *arXiv preprint arXiv:2002.02794*, 2020.