ECE209AS (Fall 2025)

Computational Robotics

Prof. Ankur Mehta ⟨mehtank@ucla.edu⟩

Lecture 8 | Reinforcement learning

Addendum to lecture videos

Errata

At the end of lec08b, I describe a modification to the Q-learning algorithm where we have one of the Q terms in the loss function update a lot slower than the other one. That is, we do gradient descent on Q_θ while leaving $\hat{Q}$ fixed, then every so often update $\hat{Q} \leftarrow Q_\theta$. I accidentally swapped which was which in the lecture; the correct assignment of Q_θ and $\hat{Q}$ is what is in the pset, reproduced here:

$$\mathcal{L}(\theta) = \| r + \gamma \max_{a'} \hat{Q}(s', a') - Q_\theta(s, a) \|$$