ECE209AS (Fall 2025)

Computational Robotics

Prof. Ankur Mehta ⟨mehtank@ucla.edu⟩

Problem set 2 | Planning / control on MDPs

Due Tue, Oct 7, 2025 @ 9am PT

Key takeaways

After this lecture, you should understand:

- How to mathematically formulate the concepts of "problem" (reward R) and "solution" (policy π) for planning and control on discrete space systems.
- The definition and purpose of value functions (V, Q), and how it relates to and differs from reward R.
- How the properties of the value functions lead to dynamic programming algorithms that exactly solve Markov Decision Problems (MDPs).

Assignment

2(a). For what systems / problems might one want a stochastic policy as opposed to a deterministic policy? In two or three English sentences with no math, describe an example of such a problem and outline why a stochastic policy would be preferred.

2(b). In one equation, write out the complete mathematical (non-recursive) definition of the value function V^π under a stochastic policy $\pi : S \times A \to \mathbb{R}$:

$$\pi(s, a) \equiv pr(a|s).$$

2(c). How might you adapt the algorithm or output of solving for an optimal (deterministic) policy such that you can instead get a stochastic policy? In one or two English sentences, explain conceptually what change you could make, and how it would address the needs you describe in 2(a) above.

2(d). Would you be willing to let us use your correct responses as (anonymized) examples for the class?