B9120-001 Dynamic Programming Fall 2023

Prof. Daniel Russo

Homework Assignment 4: Due Thursday October 5

Termination States and Discounted Problems

Do problem 1.2. of Bertsekas Volume II.

This problem shows that some problems, despite being formulated with an undiscounted objective, have an intrinsic form of discounting because any action could result in a termination state being reached. Imagine interacing with a customer who may leave the system at any time. Stochastic shortest path problems, covered in Chapter 3 of the textbook, substantially generalize this idea.

Stationary policies in finite horizon problems

Suppose there is a constant $M < \infty$ such that $-M \leq g(x,u,w) \leq M$ for every x, u, and feasible realization of the disturbance w. Consider the infinite horizon discounted formulation covered in class. Let $J^* = TJ^*$ be the optimal cost-to-go function and take $\mu^* \in G(J^*)$. Show that the policy $\pi_\mu = (\mu, \mu, \dots, \mu)$ is near optimal in a N period problem, with large but finite N, by bounding

$$\mathbb{E}^{\pi_\mu}\left[\sum_{k=0}^{N-1} \alpha^k g(x_k, u_k, w_k) \mid x_k = x\right] - \inf_{\pi=(\mu_0,\dots,\mu_{N-1})} \mathbb{E}^{\pi}\left[\sum_{k=0}^{N-1} \alpha^k g(x_k, u_k, w_k) \mid x_k = x\right].$$

Bonus Problem: Gradient Descent as a Fixed Point Iteration

You do not need to submit this problem.

Consider the simple optimization problem

$$\min_{x\in\mathbb{R}^n} \frac{1}{2} x^\top Q x - b^\top x,$$

where Q is symmetric and positive definite. Let $f(x) = \frac{1}{2}x^\top Qx - b^\top x$ and consider gradient descent with constant stepsize $\alpha > 0$:

$$x_{k+1} = x_k - \alpha \nabla f(x_k) \quad k = 0, 1, 2, \dots$$

Define the corresponding operator $H : \mathbb{R}^n \to \mathbb{R}^n$ by $H(x) = x - \alpha\nabla f(x)$ (. . . i.e. $H = I - \alpha\nabla f$). We recast the problem of minimizing f as that of finding a fixed point of H. Show that if α is chosen to be sufficiently small, then H is a contraction mapping with respect to the Euclidean norm $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$.