1. Dynamic Programming

A decision problem is characterized by (state, action, noise) \((x_k, u_k, w_k)\) for each positive integer \(k\). The dynamics is governed by: 
\[ x_{k+1} = f_k(x_k, u_k, w_k), \]
where \(f_k\) is some bounded function. Consider \(N\) (a positive integer) to be the horizon and the controller wants to minimize a cost function \(g_k(x_k, u_k, w_k)\), which is additive over the discrete time step. The terminal cost \(g_N(x_N)\) is given. Formulate the problem as a dynamic program and find the total expected cost. Define a policy sequence \(\mu = (\mu_0, \ldots, \mu_{N-1})\) as a mapping from state space to action space, i.e., \(u_k = \mu_k(x_k)\). Find optimal policy as a function of total expected cost. Also state the dynamic programming update equations for the \(k\)-th iteration, using the principle of optimality.

2. Infinite-Horizon Discounted Cost MDP

Consider a relay placement problem, where a deployment agent starts walking from state 0 on a line. He stops at regular intervals (consider the step length to be \(\delta > 0\)), and decides whether to place a relay there or not. At each step, he measures the power required to maintain a reasonable quality link. The goal of the agent is to minimize a linear combination of power cost and relay cost. Assume that the process restarts every time the agent deploys a relay. Also assume the length of the line is geometric with parameter \(\theta\). Formulate the problem as an infinite horizon MDP, with state space \((r, \gamma)\) where \(r\) and \(\gamma\) are the respective distance and power from the previously placed relay. Using the Bellman equation, find the optimal policy structure. Mention a way to compute the optimal policy.