Dynamic Programming

Aspiration behind dynamic Programming

1. Optimizing multi stage decision processes.
2. Reducing computational complexity
3. Applicability for diverse fields
4. Need to handle for deterministic and Stochastic problems

PRINCIPLE OF OPTIMALITY. An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decisions.

A sequence of decisions will be called a policy, and a policy which is most advantageous according to some preassigned criterion will be called an optimal policy.

Mathematical formulation

Discrete dynamic case

A **discrete deterministic process** involves a system that evolves over a fixed number of stages (N) through decisions that deterministically transform the system's state

The system's state is represented by a state vector p = (p1 , p2 , p3 , p4)

Each transformation Tk(p) maps a state $p \in D$ to a new state $Tk(p) \in D$.

The goal is to maximize a scalar function R(PN)

The policy is a sequence of transformations. We aim to maximize R(PN).

To maximize the overall N-stage return, we must choose the transformation Tk that maximizes the return from the remaining $N-1$ stages.

Mathematical formulation—II: Discrete stochastic case.

In a **discrete stochastic process**, the system evolves over N stages, but unlike the deterministic case, the next state after a decision (transformation) is not fixed—it follows a probability distribution.

State Vector: The system's state at any time is represented by a vector p , constrained to a region D

unlike the deterministic case, applying Tk to state p does not yield a unique next state. Instead, it produces a stochastic vector z (the new state) with an associated **vector distribution function** dGk(p,z).

The return R(z) depends on the final state z after N stages. However, due to stochasticity, maximizing R(z) directly is meaningless because z is random. Instead, the value of a policy is measured by the **expected return**, defined as the average value of R(z) over all possible final states weighted by their probabilities.

Expected Return: For a given Tk, the next state z is random with distribution dGk(p,z). The return from the last N−1 stages starting from z is fN−1(z) , and the expected return over all possible z is:

$$\int_{z\in D} f_{N-1}(z)\, dG_k(p, z)$$

Mathematical Formulation—III: Infinite Stochastic Process

The goal of infinite stochastic process is to maximize the expected return over infinite input.

$$f(p) = \max_{k} \int_{s \in D} f(s)\, dG_k(p, s),$$

6. Mathematical formulation—IV: Continuous deterministic process

controlling a system (like a car's position x(t)) over time [0,T] by choosing a control u(t) (like speed). The state changes via dx/dt=G(x,u) , and you aim to maximize a total reward, like ∫F(x,u) dt ,(e.g., minimizing fuel use).

As S→0, the equation becomes a partial differential equation (PDE):

$$\frac{\partial f}{\partial T} = \max_u \left[F(p, u) + G(p, u)\frac{\partial f}{\partial p}\right],$$