Iterative Algorithm for Computing the Value of a Markov Reward Process (MRP)

We want to compute the value function $V(s)$ for each state $s \in S$ in a Markov Reward Process. The value function satisfies the Bellman equation:

$$V(s) = R(s) + \gamma \sum_{s' \in S} P(s'|s)V(s'),$$

where

- $R(s)$ is the expected immediate reward at state s,
- $P(s'|s)$ is the transition probability from s to s',
- $\gamma \in [0, 1)$ is the discount factor.

Iterative Algorithm (Dynamic Programming)

1. **Initialize:**
$$V_0(s) = 0 \quad \text{for all } s \in S.$$

2. **Iterate:** for $k = 1, 2, \ldots,$
$$V_k(s) = R(s) + \gamma \sum_{s' \in S} P(s'|s)V_{k-1}(s'), \quad \forall s \in S.$$

3. **Stopping Criterion:** stop when
$$\max_{s \in S} \left|V_k(s) - V_{k-1}(s)\right| < \epsilon,$$
where $\epsilon > 0$ is a small tolerance (e.g., 10^{-6}).

Computational Complexity

Each iteration requires updating all states, which involves summing over all possible next states. Therefore, the complexity is

$$O(|S|^2) \quad \text{per iteration, where } |S| = N.$$