MDP - 1d

$T = \{0, 1, 2, 3, 4, 5\}$. ⟶ time space

$S = \{-3, \ldots, 3\}$ ⟶ state space

$A = \{-1, 1\}$. ⟶ action space

Trans. prob

act \ Prob	DN	Flat	up
1		1/2	1/2
−1	1/2	1/2	

DN → flat if state $= -3$

Up → flat if state $= 3$

Ter. cost

$$g(x) = x^2$$

obj min terminal cost.

soln

① opt policy $= \begin{cases} 1 & \text{if } x < 0 \\ -1 & \text{if } x > 0 \end{cases}$

② s-val