

Lecture Notes 3

Minfei Chen

2025 年 8 月 14 日

1 最优策略 (optimal policy)

1.1 定义

一个策略 π^* 是最优策略，如果 $v_{\pi^*}(s) \geq v_{\pi}(s)$ 对于所有状态 s 和其他所有策略 π 。

1.2 贝尔曼最优公式 (BOE, elementwise form)

$$v(s) = \max_{\pi} \sum_a \pi(a|s) \left(\sum_r p(r|s, a) r + \gamma \sum_{s'} p(s'|s, a) v(s') \right), \quad \forall s \in \mathcal{S}$$

* 已知：系统信息（状态转移概率、reward、 γ ）；求解：最优策略 π