

证明

16337341 朱志儒

用线性特征函数作为 Q 函数的近似函数：

$$Q^*(s, a) \approx Q(s, a, w) = w_1 f_1(s, a) + w_2 f_2(s, a) + \cdots + w_n f_n(s, a)$$

最小化损失函数：

$$J(w) = E \left[(Q^*(s, a) - Q(s, a, w))^2 \right]$$

使用梯度下降法寻找局部最优解，即对上式两边同时对 w_i 求偏导，得到修正量 Δw_i ：

$$\begin{aligned} \frac{\partial J(w)}{\partial w_i} &= -2(Q^*(s, a) - Q(s, a, w)) \frac{\partial Q(s, a, w)}{\partial w_i} \\ \Delta w_i &= -2(Q^*(s, a) - Q(s, a, w)) f_i(s, a) \\ \Delta w_i &= \alpha [\text{error}] f_i(s, a) \end{aligned}$$

所以 w_i 的迭代公式：

$$w_i \leftarrow w_i + \alpha [\text{error}] f_i(s, a)$$