

9-8

$$\begin{aligned} E &= [U_t - \hat{v}(s_t, a_t)]^2 \\ &\rightarrow -U_t \ln \hat{v}(s_t, w_t) \end{aligned}$$

$$\begin{aligned} w_{t+1} &= w_t - \alpha \nabla E \\ &= w_t + \alpha \frac{U_t}{\hat{v}(s_t, a_t)} \nabla \hat{v}(s_t, w_t) \\ &= w_t + \alpha \frac{U_t}{\hat{v}(s_t, a_t)} x \end{aligned} \quad (\text{when linear})$$