Lecture-14

- **Function approximation**

Q: How can we estimate V^π or Q^π if the states (or actions) are continuous

i.e. $s_t \in \mathbb{R}$ or $s_t \in \mathbb{R}^n$

- Regression problem

① Use parameterized function

$V_\omega(s)$

... where ω = weight vector $\omega \in \mathbb{R}^m$

e.g. Linear function approximations

$$V_\omega(s) \triangleq \omega^T \phi(s),$$

... where $\phi(s) \in \mathbb{R}^m$ is called feature vector

- Polynomial regression

$$\phi(s) = \begin{bmatrix} 1 \\ s \\ \vdots \\ s^{m-1} \end{bmatrix} \qquad V_\omega(s) = \omega_1 + \omega_2 s + \omega_3 s^2 + \dots$$

e.g. in case of fourier basis

$$\phi(s) = \begin{bmatrix} \cos(0) \\ \cos(\pi s) \\ \cos(2\pi s) \\ \cos(3\pi s) \end{bmatrix} \quad \text{assuming } s \in [0, 1]$$

② Tabular

$$\phi(s) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \boxed{1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \text{at } s^{th} \text{ place}$$

... where, $s \in \{1, \cdots, m\}$

③ Neural Network

$V_w(s)$ = output of a neural network with weights w, and inputs 's'.

Benefits & Drawbacks :

Benefits - ① Generalization - can estimate $V^{\pi}(s)$, even if we have seen ~~~~ state s_1 (very close to s) but not s.

Benefits - ② Can handle continuous states

③ Will allow us to handle partial observability

④ Do not assume that $\phi(s)$ is a Markovian state construction.

$$\phi(x, y) = X$$

Drawbacks - ① V^{π} may not be representable

② Fewer convergence guarantees

Derivation:

$$L = \mathbb{E}\left[\frac{1}{2}\left(V^{\pi}(s) - V_w(s_t)\right)^2\right]$$

$$L = \mathbb{E}\left[\frac{1}{2}\left(\left(R_t + \gamma \cdot V_w(s_{t+1})\right) - V_w(s_t)\right)^2\right]$$

$$L = \mathbb{E}\left[\frac{1}{2}\delta^2\right]$$

$$w \leftarrow w - \alpha \cdot \frac{\partial L(w)}{\partial w}$$

$$\therefore \frac{\partial L(\omega)}{\partial \omega} = \mathbb{E}\left[(R_t + \gamma \cdot V_\omega(S_{t+1}) - V_\omega(S_t)) \right]$$

$$\times \frac{\partial}{\partial \omega} (R_t + \gamma \cdot V_\omega(S_{t+1}) - V_\omega(S_t))$$

$$= \mathbb{E}[\delta_t] \times \left(\gamma \frac{\partial V_\omega(S_{t+1})}{\partial \omega} - \frac{\partial V_\omega(S_t)}{\partial \omega} \right)$$

$$\therefore \omega \leftarrow \omega - \alpha \cdot \delta_t \left(\gamma \frac{\partial V_\omega(S_{t+1})}{\partial \omega} - \frac{\partial V_\omega(S_t)}{\partial \omega} \right)$$

$$\therefore \omega \leftarrow \omega + \alpha \cdot \delta_t \left(\frac{\partial V_\omega(S_t)}{\partial \omega} - \gamma \cdot \frac{\partial V_\omega(S_{t+1})}{\partial \omega} \right)$$

→ What direction of change to ω will most quickly increase $V_\omega(S_t)$

→ This term reduces $V(S_{t+1})$, ~~the~~ – slows down process (ignore the term)

$\therefore$ TD – update with function approximat^n

$$\omega \leftarrow \omega + \alpha \cdot \delta_t \cdot \frac{\partial V_\omega(S_t)}{\partial \omega}$$

- tabular

$$V_w(s) = w^T \phi(s)$$

$$\frac{\partial V_w(s)}{\partial w} = \phi(s) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

$$w \leftarrow w + \alpha \cdot \delta_t \cdot \phi(s)$$

$$V(s_t) \leftarrow V(s_t) + \alpha \cdot \delta_t$$

Properties -

Tabular - converges to V^π almost surely if α is decayed properly

- Converges to V^π in the mean if α is a sufficiently small constant

Linear - Guaranteed to converge to some w_∞, such that

$$MSE(w_\infty) \leq \frac{1}{1-\lambda} MSE(w^*)$$