Exercise 7

Kai Schneider

June 15, 2021

Task 1 Linear function approximation

a.)

With linear function approximation we have a value function with the following form:
$\hat{v}(s,w) = w^T x(s) = \sum_{i=1}^{d} w_i x_i(s)$
For tabular methods we can write all $x_i(s)$ as $x_i(s) = [0 \dots \underbrace{v(s)}_{\text{i-th entry}} \dots 0]^T$

This way our value function can be written the same as for the linear function approximating with the weight vector $w = [1 \dots 1]$.
This way the tabular version is only a special case of the linear function approximation with the $dim(w) = |S|$.

b.)

Sarsa: $Q(S_t,A_t) \leftarrow Q(S_t,A_t) + \alpha\delta_t$ with $\delta_t = [R_{t+1} + \gamma Q(S_{t+1},A_{t+1}) - Q(S_t,A_t)]$

I: tabular

For the tabular case we can directly use the formula given in the slides (v05s15):

$$Q(S_t,A_t) \leftarrow Q(S_t,A_t) + \alpha[R_{t+1} + \gamma Q(S_{t+1},A_{t+1}) - Q(S_t,A_t)]$$

II: function approximation

For state-action values we have the following update rule for weights in control:

$w_{t+1} \leftarrow w_t + \alpha[U_t - \hat{q}(S_t,A_t,w)]\nabla\hat{q}(S_t,A_t,w)$
with $U_t = R_{t+1} + \gamma\hat{q}(S_{t+1},A_{t+1},w)$ for one-step SARSA

plugged in we get: $w_{t+1} \leftarrow w_t + \alpha[R_{t+1} + \gamma\hat{q}(S_{t+1},A_{t+1},w) - \hat{q}(S_t,A_t,w)]\nabla\hat{q}(S_t,A_t,w)$

III: linear function approximation

With the value function beeing $\hat{v}(s,w) = w^T x(s)$ for the linear case and the update formula for the linear TD $w \leftarrow w + \alpha[R_{t+1} + \gamma w^T x(S_{t+1}) - w^T x(S_t)]x(S_t)$ we can get:

$$w_{t+1} \leftarrow w_t + \alpha[R_{t+1} + \gamma w^T x(S_{t+1},A_{t+1}) - w^T x(S_t,A_t)]x(S_t,A_t)$$