Stanford CME 241 (Winter 2021) - Assignment 12

Assignments:

1. Implement the n-Step Bootstrapping Prediction algorithm from scratch in Python code. First do it for the Tabular case, then do it for the case of Function Approximation.

2. Implement the TD(λ) Prediction algorithm from scratch in Python code. First do it for the Tabular case, then do it for the case of Function Approximation.

3. Prove that the MC Error can be written as the sum of discounted TD errors, i.e.,

$$G_t - V(S_t) = \sum_{u=t}^{T-1} \gamma^{u-t} \cdot (R_{u+1} + \gamma \cdot V(S_{u+1}) - V(S_u))$$

   The goal here is for you to practice formal proof-writing of these types of simple yet important identities. So aim to work this out from scratch rather than treating this as a special case of a more general result proved in class or in the textbook.

4. Test your above implementation of TD(λ) Prediction algorithm by comparing the Value Function of an MRP you have previously developed (or worked with) as obtained by Policy Evaluation (DP) algorithm, as obtained by MC, as obtained by TD, and as obtained by your TD(λ) implementation. Plot graphs of convergence for different values of λ.

3. Prove that the MC Error can be written as the sum of discounted TD errors, i.e.,

$$G_t - V(S_t) = \sum_{u=t}^{T-1} \gamma^{u-t} \cdot (R_{u+1} + \gamma \cdot V(S_{u+1}) - V(S_u))$$

The goal here is for you to practice formal proof-writing of these types of simple yet important identities. So aim to work this out from scratch rather than treating this as a special case of a more general result proved in class or in the textbook.

$$G_t = \sum_{i=t+1}^{T} \gamma^{i-t-1} R_i = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots + \gamma^{T-t-1} R_T = R_{t+1} + \gamma G_{t+1}$$

$$\sum_{u=t}^{T-1} \gamma^{u-t} (R_{u+1} + \gamma V(S_{u+1}) - V(S_u))$$
$$= [R_{t+1} + \gamma V(S_{t+1}) - V(S_t)] + \gamma [R_{t+2} + \gamma V(S_{t+2}) - V(S_{t+1})]$$
$$+ \cdots + \gamma^{T-1-t} [R_T + \gamma V(S_T) - V(S_{T-1})]$$

$$G_t - V(S_t) - \sum_{u=t}^{T-1} \gamma^{u-t} (R_{u+1} + \gamma V(S_{u+1}) - V(S_u))$$

$$= \cancel{R_{t+1}} + \cancel{\gamma R_{t+2}} + \cancel{\gamma^2 R_{t+3}} + \cdots + \cancel{\gamma^{T-t-1} R_T} - \cancel{V(S_t)}$$
$$- \{ [\cancel{R_{t+1}} + \gamma V(S_{t+1}) - \cancel{V(S_t)}] + \gamma [\cancel{R_{t+2}} + \gamma V(S_{t+2}) - V(S_{t+1})]$$
$$+ \cdots + \gamma^{T-1-t} [\cancel{R_T} + \gamma V(S_T) - V(S_{T-1})] \}$$

$$= \cancel{\gamma V(S_{t+1})} - \cancel{\gamma V(S_{t+1})} + \cancel{\gamma^2 V(S_{t+2})} \cdots - \cancel{\gamma^{T-1-t} V(S_{T-1})} + \gamma^{T-t} V(S_T)$$
$$= 0$$