Attempts to exercise in Reinforcement Learning book Chapter 7

Mengliao Wang

August 12, 2017

Exercise 7.1:

First of all, according to definition of G_t and $G_{t:t+n}$ we have the following equations:

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + ... + \gamma^{T-t-1} R_T \\ G_{t+n} &= R_{t+n+1} + \gamma R_{t+n+2} + ... + \gamma^{T-t-n-1} R_T \\ G_{t:t+n} &= R_{t+1} + \gamma R_{t+2} + ... + \gamma^{n-1} R_{t+n} + \gamma^n V_{t+n-1}(S_{t+n}) \end{aligned}$$

From these equation we can have

$$G_t = G_{t:t+n} - \gamma^n V_{t+n-1}(S_{t+n}) + \gamma^n G_{t+n} \tag{1}$$

Then by applying equation 1, the difference between G_t and $V_{t+n-1}(S_t)$ can be written as:

$$\begin{aligned} G_t - V_{t+n-1}(S_t) &= G_{t:t+n} - \gamma^n V_{t+n-1}(S_{t+n}) + \gamma^n G_{t+n} - V_{t+n-1}(S_t) && (2) \\ &= [G_{t:t+n} - V_{t+n-1}(S_t)] + \gamma^n [G_{t+n} - V_{t+n-1}(S_{t+n})] && (3) \\ &= \delta_t^n + \gamma^n [\delta_{t+n}^n + \gamma^n [G_{t+2n} - V_{t+2n-1}(S_{t+2n})]] && (4) \\ &= \sum_{k=0}^{(T-t)/n} \gamma^{kn} \delta_{t+kn}^n + r && (5) \end{aligned}$$

Here r is the residual if $T - t - 1$ cannot be divided by n without remainder. Assuming the last $t + nk$ time step before reaching T is $t + nK$, then we can write r as:

$$\begin{aligned} r &= \gamma^{t+nK} [G_{t+nK} - V_{t+n-1}(S_{t+nK}] && (6) \\ & && (7) \end{aligned}$$