Mathematical Foundations of Reinforcement Learning

Homework 2

Due Date*: May 08, 2017

Consider multistep simulation-based methods to solve

$$\Phi r = \Pi T^{(\lambda)} \Phi r \tag{1}$$

where $T^{(i)} = (1-\lambda) \sum_{l=0}^{\infty} \lambda^l T^{l+1}$

Problem 1. Show that (6.92) and (6.93) hold

Problem 2. Show that (6.94), (6.95) and (6.96) hold

Problem 3. Show that

$$\begin{aligned} Z_k &= \alpha\lambda Z_{k-1} + \phi(i_k) \\ C_k^{(\lambda)} &= (1-\delta_k) C_{k-1}^{(\lambda)} + \delta_k Z_k (\phi(i_k) - \alpha\phi(i_{k+1})' \\ d_k^{(i)} &= (1-\delta_k) d_{k-1}^{(\lambda)} + \delta_k Z_k g(i_k, i_{k+1}) \\ \delta_k &= \frac{1}{k+1} \end{aligned} \tag{2}$$

with $Z_{-1} = 0, C_{-1} = 0, d_{-1} = 0$

Problem 4. Show that $LSPE(\lambda)$ method can be written in the following least square form

$$r_{k+1} = \arg\min_{r \in R^s} \sum_{t=0}^{K} (\phi(i_t)' r - \phi(i_t)' r_k - \sum_{m=t}^{K} (\alpha\lambda)^{m-t} d_k(i_m, i_{m+1}))^2 \tag{3}$$

where the temporal differences

$$d_k(i_t, i_{t+1}) = g(i_t, t_{t+1}) + \alpha\phi(i_{t+1})' r_k - \phi(i_t)' r_k \tag{4}$$