V and Q recursions

$$V^{\pi}(s) = \mathbb{E}_{\pi}\left\{\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+1} | s_t = s\right\} \tag{1}$$

$$Q^{\pi}(s,a) = \mathbb{E}_{\pi}\left\{\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+1} | s_t = s, a_t = a\right\} \tag{2}$$

From Sutton's book page 70, Equation (3.10), recursive formulation of the value function update is:

$$V^{\pi}(s) = \sum_{a}\pi(s,a)\underbrace{\sum_{s'}P^{a}_{ss'}\left[R^{a}_{ss'} + \gamma V^{\pi}(s')\right]}_{Q^{\pi}(s,a)} \tag{3}$$

$$= \sum_{a}\pi(s,a)Q^{\pi}(s,a) \tag{4}$$

As an exercise (Exercise 3.8, page 72 in Suttons book), the same recursion has to be derived, but this time for the Q-Value function, Equation 2.

$$\begin{aligned}
Q^{\pi}(s,a) &= \mathbb{E}_{\pi}\left\{\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+1} | s_t = s, a_t = a\right\} \\
&= \mathbb{E}_{\pi}\left\{r_{t+1} + \gamma\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+2} | s_t = s, a_t = a\right\} \\
&= \mathbb{E}_{\pi}\left\{r_{t+1} | s_t = s, a_t = a\right\} + \gamma\,\mathbb{E}_{\pi}\left\{\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+1} | s_t = s, a_t = a\right\} \\
&= \sum_{s'}P^{a}_{ss'}R^{a}_{ss'} + \gamma\sum_{s'}P^{a}_{ss'}\,\mathbb{E}_{\pi}\left\{\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+2} | s_{t+1} = s', a_t = a\right\} \\
&= \sum_{s'}P^{a}_{ss'}R^{a}_{ss'} + \gamma\sum_{s'}P^{a}_{ss'}\sum_{a'}\pi(s',a')\,\mathbb{E}_{\pi}\left\{\sum_{k=0}^{\infty}\gamma^{k} r_{t+k+2} | s_{t+1} = s', a_{t+1} = a'\right\} \\
&= \sum_{s'}P^{a}_{ss'}R^{a}_{ss'} + \gamma\sum_{s'}P^{a}_{ss'}\sum_{a'}\pi(s',a')Q(s',a') \\
&= \sum_{s'}P^{a}_{ss'}\Big[R^{a}_{ss'} + \gamma\underbrace{\sum_{a'}\pi(s',a')Q(s',a')}_{V^{\pi}(s')}\Big]
\end{aligned}$$

We not that the last term in the expansion of Q, contains Equation 4. So we can substitute it back in and get:

$$Q^{\pi}(s,a) = \sum_{s'}P^{a}_{ss'}\Big[R^{a}_{ss'} + \gamma V^{\pi}(s')\Big] \tag{5}$$