① ASSIGNMENT 3

ALEC PORTELLI

①

Ⓐ

Ⓘ

We can rewrite $Q_{n+1} = Q_n + \alpha_n (G_n - Q_n)$ as the following:

$$Q_{n+1} = Q_n + \frac{1}{n}(G_n - Q_n)$$

Using induction, the base case becomes:

$$Q_{1+1} = Q_1 + \frac{1}{1}(G_1 - Q_1) \text{ where } Q_1 = 0$$

$$Q_2 = 0 + 1(G_1 - 0)$$

$$Q_2 = G_1$$

Induction step then becomes: $Q_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} G_i$

We can rewrite as: $Q_{n+1} = \frac{n-1}{n} Q_n + \frac{1}{n} G_n$

Substitute: $Q_{n+1} = \frac{n-1}{n} \cdot \frac{1}{n} \sum_{i=1}^{n} G_i + \frac{1}{n} G_n$

Multiply: $Q_{n+1} = \frac{n-1}{n^2} \sum_{i=1}^{n} G_i + \frac{1}{n} G_n$

Multiply: $Q_{n+1} = \frac{1}{n+1}\left(\frac{n}{n-1} \cdot \frac{n-1}{n} \sum_{i=1}^{n} G_i + G_n\right)$

Simplify: $Q_{n+1} = \frac{1}{n+1}\left(\frac{n}{n-1} \sum_{i=1}^{n} G_i + G_n\right)$

We can remove the $\frac{n}{n-1}$ because it is constant, so $Q_{n+1} = \frac{1}{n} \sum_{i=1}^{n} G_i + G_n$

BACK

②

(A) (II)

$$V_n = E\left[(Q_n - r)^2\right] = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} G_i - r\right)^2\right]$$ which can be rewritten as:

$$V_n = \frac{1}{n^2}\sum_{i=1}^{n} E\left[(G_i - r)^2\right]$$

We can swap out $E\left[(G_i-r)^2\right]$ as the variance of G_i, which yields:

$$V_n = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}[G_i]$$

Take the limit: $\lim_{n\to\infty} V_n \; \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}[G_i]$ which is the same as: $\lim_{n\to\infty} V_n = \phi$!!

(B) (I)

$$V_{n+1} = E\left[(Q_{n+1} - r)^2\right] = E\left[(Q_n + \alpha(G_n - G_n) - r)^2\right]$$ which expands to:

$$V_{n+1} = E\left[(1-\alpha)^2 Q_n^2 + 2\alpha(1-\alpha)(G_n - r)Q_n + \alpha^2(G_n - r)^2\right]$$ which

can be rewritten as: $V_{n+1} = (1-\alpha)^2 E[Q_n^2] + 2\alpha(1-\alpha)E[(G_n - r)G_n] + \alpha^2 E[(G_n - r)^2]$

Because r is constant we simplify: $V_{n+1} = (1-\alpha)^2 E[G_n^2] + \alpha^2 E[(G_n - r)^2]$

We can substitute back in for variance $\therefore V_{n+1} = (1-\alpha)^2 V_n + \alpha^2 \mathrm{Var}(G_n)$

(II)

We can rewrite $\lim_{n\to\infty}\left[V_{n+1} - \alpha \mathrm{Var}[G_n]\right] = \lim_{n\to\infty}\left[(1-\alpha)^2 V_n\right]$

$\lim_{n\to\infty}\left[(1-\alpha)^2 V_n\right] = (1-\alpha)^2 \lim_{n\to\infty} V_n$

$\lim_{n\to\infty} V_n = 0$, we can substitute and rewrite as: $\lim_{n\to\infty}\left[(1-\alpha)^2 V_n\right] = (1-\alpha^2)\cdot 0 = 0$

$\therefore \lim_{n\to\infty}\left[(1-\alpha)^2 V_n\right] = (1-\alpha)^2 \cdot \emptyset = \emptyset$

(2)

Sample averages are used for action values estimation because sample averages <u>do not</u> yield bias that constant step sizes do.

Sample averages ~~would~~ <u>do not</u> perform well in non stationary problems, so we can change the step size

We can see that Q_1 is weighted by $\prod_{i=1}^{n}(1-\alpha_i)$, so when $i=1$, that means $\beta_n = \alpha$ and that the weighted step size $\to 0 \; \forall i$. This means that Q_i is no longer a factor in the calculation of Q_{n+1}

We can see from the mathematical deduction:

$\bar{o}_0 = 0$ and $\bar{o}_1 = \alpha$

$\bar{o}_{n+1} = \bar{o}_n + \alpha(1-\bar{o}_n) = \alpha + (1-\alpha)\bar{o}_n$

$\alpha + \alpha(1-\alpha) + \alpha(1-\alpha)^2 + \alpha(1-\alpha)^3 \bar{o}_{n-2}$

which can be simplified to: $\alpha \sum_{i=0}^{n-1} (1-\alpha)^i$

$$\beta_n = \frac{\alpha}{\bar{o}_n} = \frac{\alpha}{\alpha \sum_{i=0}^{n-2} (1-\alpha)^i} = \frac{1}{\sum_{i=0}^{n-2} (1-\alpha)^i}$$