20173261 배정민.

(i)

Prob 1) $C_k^{(\lambda)} = \frac{1}{k+1}\sum_{t=0}^{k} \phi(i_t)\boxed{\sum_{m=t}^{k}\alpha^{m-t}\lambda^{m-t}(\phi(i_m)-\alpha\phi(i_{m+1}))'}$

$$\sum_{m=t}^{k}\alpha^{m-t}\lambda^{m-t}(\phi(i_m)-\alpha\phi(i_{m+1}))'$$

$$= \phi(i_t)-\alpha\phi(i_{t+1})$$
$$\alpha\lambda\phi(i_{t+1})-\alpha^2\lambda\phi(i_{t+2})$$
$$\alpha^2\lambda^2\phi(i_{t+2})-\alpha^3\lambda^2\phi(i_{t+3})$$
$$\vdots$$
$$+\ \alpha^{k-t}\lambda^{k-t}\phi(i_k)-\alpha^{k-t+1}\lambda^{k-t}\phi(i_{k+1})$$

$$\phi(i_t)-\alpha(1-\lambda)\phi(i_{t+1})-\alpha^2\lambda\overset{(1-\lambda)}{\wedge}\phi(i_{t+2})+\cdots+\alpha^{k-t}\lambda^{k-t-1}(1-\lambda)\phi(i_k)-\alpha^{k-t+1}\lambda^{k-t}\phi(i_{k+1})$$

$$=\phi(i_t)-\alpha(1-\lambda)\sum_{m=t}^{k-1}\alpha^{m-t}\lambda^{m-t}\phi(i_{m+1})-\cancel{\alpha^{k-t+1}\lambda^{k-t}\phi(i_{k+1})}$$

$$=\phi(i_t)-\alpha(1-\lambda)\sum_{m=t}^{k-1}\alpha^{m-t}\lambda^{m-t}\phi(i_{m+1})$$

$$\therefore C_k^{(\lambda)}=\frac{1}{k+1}\sum_{t=0}^{k}\phi(i_t)\left(\phi(i_t)-\underline{\alpha(1-\lambda)\sum_{m=t}^{k-1}\alpha^{m-t}\lambda^{m-t}\phi(i_{m+1})}\right)'$$

이때

$$C^{(\lambda)}=\Phi'\Xi(I-\alpha P^{(\lambda)})\Phi=\sum_{i=1}^{n}\xi_i\phi(i)\left(\phi(i)-\alpha\sum_{j=1}^{n}p_{ij}^{(\lambda)}\phi(j)\right)'$$ 이고

$$\frac{1}{k+1}\sum_{t=0}^{k}\phi(i_t)\phi(i_t)'\longrightarrow\sum_{i=1}^{n}\xi_i\phi(i)\phi'(i)$$

이고

$$p_{ij}^{(\lambda)}=\underbrace{(1-\lambda)\sum_{l=0}^{\infty}\alpha^l\lambda^l p_{ij}^{(l+1)}}, \left(P^{(\lambda)}=(1-\lambda)\sum_{l=0}^{\infty}\alpha^l\lambda^l P^{l+1}\right)$$

이므로

$$\frac{1}{k+1}\sum_{t=0}^{k}\phi(i_t)\left((1-\lambda)\sum_{m=t}^{k-1}\alpha^{m-t}\lambda^{m-t}\phi(i_{m+1})\right)'\rightarrow\sum_{i=1}^{n}\xi_i\phi(i)\sum_{j=1}^{n}p_{ij}^{(\lambda)}\phi(j)'$$

$$\therefore C_k^{(\lambda)}\rightarrow\sum_{i=1}^{n}\xi_i\phi(i)\left(\phi(i)-\alpha\sum_{j=1}^{n}p_{ij}^{(\lambda)}\phi(j)\right)'=C^{(\lambda)}$$

(ii) 같은 방법으로,

$$d_k^{(\lambda)}=\frac{1}{k+1}\sum_{t=0}^{k}\phi(i_t)\sum_{m=t}^{k}\alpha^{m-t}\lambda^{m-t}g_{i_m}$$

$$d^{(\lambda)}=\sum_{i=1}^{n}\xi_i\phi(i)g_i^{(\lambda)}.\qquad g_i^{(\lambda)}=\sum_{l=0}^{\infty}\alpha^l\lambda^l P^l g_i.$$

$$\frac{1}{k+1}\sum_{t=0}^{k}\phi(i_t)\sum_{m=t}^{k}\alpha^{m-t}\lambda^{m-t}g_{i_m}\longrightarrow\sum_{i=1}^{n}\xi_i\phi(i)g_i^{(\lambda)}=d^{(\lambda)}\ //.$$

Prob 2) $Z_t = \sum_{m=0}^{t} (\alpha\lambda)^{t-m} \phi(\bar{i}_m)$

(i) $(k+1)C_k = \sum_{t=0}^{k} \phi(\bar{i}_t) \sum_{m=t}^{k} \alpha^{m-t}\lambda^{m-t} (\phi(\bar{i}_m) - \alpha\phi(\bar{i}_{m+1}))'$

$$= \sum_{m=0}^{k} (\alpha\lambda)^m \phi(\bar{i}_0) (\phi(\bar{i}_m) - \alpha\phi(\bar{i}_{m+1}))$$

$$+ \sum_{m=1}^{k} (\alpha\lambda)^{m-1} \phi(\bar{i}_1) (\phi(\bar{i}_m) - \alpha\phi(\bar{i}_{m+1}))$$

$$+ \sum_{m=k-1}^{k} (\alpha\lambda)^{m-k+1} \phi(\bar{i}_{k-1}) (\phi(\bar{i}_m) - \alpha\phi(\bar{i}_{m+1}))$$

$$+ \sum_{m=k}^{k} (\alpha\lambda)^{m-k} \phi(\bar{i}_k) (\phi(\bar{i}_m) - \alpha\phi(\bar{i}_{m+1}))$$

$$= (\phi(\bar{i}_0) - \alpha\phi(\bar{i}_1))\phi(\bar{i}_0) +$$

$$(\phi(\bar{i}_1) - \alpha\phi(\bar{i}_2)) \sum_{m=0}^{1} (\alpha\lambda)^{1-m}\phi(\bar{i}_m) +$$

$$(\phi(\bar{i}_2) - \alpha\phi(\bar{i}_3)) \sum_{m=0}^{2} (\alpha\lambda)^{2-m}\phi(\bar{i}_m) +$$

$$\vdots$$

$$(\phi(\bar{i}_k) - \alpha\phi(\bar{i}_{k+1})) \sum_{m=0}^{k} (\alpha\lambda)^{k-m}\phi(\bar{i}_m)$$

$$= \sum_{t=0}^{k} (\phi(\bar{i}_t) - \alpha\phi(\bar{i}_{t+1})) \sum_{m=0}^{t} (\alpha\lambda)^{t-m}\phi(\bar{i}_m)$$

$$= \sum_{t=0}^{k} \sum_{m=0}^{t} \underline{(\alpha\lambda)^{t-m}\phi(\bar{i}_m)} (\phi(\bar{i}_t) - \alpha\phi(\bar{i}_{t+1}))$$

$$= \sum_{t=0}^{k} Z_t (\phi(\bar{i}_t) - \alpha\phi(\bar{i}_{t+1}))'$$

(ii) $(k+1)d_k = \sum_{t=0}^{k} \phi(\bar{i}_t) \sum_{m=t}^{k} \alpha^{m-t} \lambda^{m-t} g_{\bar{i}_m}$

$$= \sum_{m=0}^{k} (\alpha\lambda)^m \phi(\bar{i}_0) g_{\bar{i}_m}$$

$$+ \sum_{m=1}^{k} (\alpha\lambda)^{m-1}\phi(\bar{i}_1) g_{\bar{i}_m} + \cdots \sum_{m=k-1}^{k} (\alpha\lambda)^{m-k+1} \phi(\bar{i}_{k-1}) g_{\bar{i}_m} + \sum_{m=k}^{k} (\alpha\lambda)^{m-k} \phi(\bar{i}_k) g_{\bar{i}_m}$$

$$= g_{\bar{i}_0}\phi(\bar{i}_0) + g_{\bar{i}_1} \sum_{m=0}^{1} (\alpha\lambda)^{1-m}\phi(\bar{i}_m) + \cdots g_{\bar{i}_k} \sum_{m=0}^{k} (\alpha\lambda)^{k-m}\phi(\bar{i}_m)$$

$$= \sum_{t=0}^{k} g(\bar{i}_t, \bar{i}_{t+1}) \sum_{m=0}^{t} (\alpha\lambda)^{t-m}\phi(\bar{i}_m)$$

$$= \sum_{t=0}^{k} Z_t\, g(\bar{i}_t, \bar{i}_{t+1})$$

Prob 3 (i) $z_k = \sum_{m=0}^{k} (\alpha\lambda)^{k-m} \phi(\bar{i}_m)$

$$= (\alpha\lambda)^k \phi(\bar{i}_0) + (\alpha\lambda)^{k-1} \phi(\bar{i}_1) + \cdots (\alpha\lambda)^1 \phi(\bar{i}_{k-1}) + (\alpha\lambda)^0 \phi(\bar{i}_k)$$

$$z_{k-1} = \sum_{m=0}^{k-1} (\alpha\lambda)^{k-1-m} \phi(\bar{i}_m)$$

$$= (\alpha\lambda)^{k-1} \phi(\bar{i}_0) + (\alpha\lambda)^{k-2} \phi(\bar{i}_1) + \cdots + (\alpha\lambda)^0 \phi(\bar{i}_{k-1})$$

$\therefore (\alpha\lambda) z_{k-1} + \phi(\bar{i}_k)$

$$= (\alpha\lambda)^k \phi(\bar{i}_0) + (\alpha\lambda)^{k-1} \phi(\bar{i}_1) + \cdots + (\alpha\lambda)^1 \phi(\bar{i}_{k-1}) + \phi(\bar{i}_k)$$

$$= z_k \,/\!/$$

(ii) $C_k^{(\lambda)} = \frac{1}{k+1} \sum_{t=0}^{k} z_t (\phi(\bar{i}_t) - \alpha \phi(\bar{i}_{t+1}))'$

$$= \frac{1}{k+1} \Big[z_0 (\phi(\bar{i}_0) - \alpha \phi(\bar{i}_1)) + z_1 (\phi(\bar{i}_1) - \alpha \phi(\bar{i}_2)) + \ldots + \overset{z_{k-1}(\phi(\bar{i}_{k-1}) - \alpha\phi(\bar{i}_k))}{} + z_k (\phi(\bar{i}_k) - \alpha \phi(\bar{i}_{k+1})) \Big]$$

$C_{k-1}^{(\lambda)} = \frac{1}{k} \sum_{t=0}^{k-1} z_t (\phi(\bar{i}_t) - \alpha \phi(\bar{i}_{t+1}))'$

$$= \frac{1}{k} \Big[z_0 (\phi(\bar{i}_0) - \alpha \phi(\bar{i}_1)) + z_1 (\phi(\bar{i}_1) - \alpha \phi(\bar{i}_2))' + \cdots z_{k-1} (\phi(\bar{i}_{k-1}) - \alpha \phi(\bar{i}_k)) \Big]$$

$\therefore (1-\delta_k) C_{k-1}^{(\lambda)} + \delta_k z_k (\phi(\bar{i}_k) - \alpha\phi(\bar{i}_{k+1}))'$

$$= \frac{k}{k+1} C_{k-1}^{(\lambda)} + \frac{1}{k+1} z_k (\phi(\bar{i}_k) - \alpha \phi(\bar{i}_{k+1}))'$$

$$= \frac{\cancel{k}}{k+1} \times \frac{1}{\cancel{k}} \Big[z_0 (\phi(\bar{i}_0) - \alpha \phi(\bar{i}_1)) + \cdots z_{k-1} (\phi(\bar{i}_{k-1}) - \alpha \phi(\bar{i}_k)) \Big] + \frac{1}{k+1} z_k (\phi(\bar{i}_k) - \alpha \phi(\bar{i}_{k+1}))$$

$$= \frac{1}{k+1} \Big[z_0 (\phi(\bar{i}_0) - \alpha \phi(\bar{i}_1)) + \cdots z_k (\phi(\bar{i}_k) - \alpha \phi(\bar{i}_{k+1})) \Big] = C_k^{(\lambda)} \,/\!/$$

(iii) $d_k^{(\lambda)} = \frac{1}{k+1} \sum_{t=0}^{k} z_t \, g(\bar{i}_t, \bar{i}_{t+1})$

$$= \frac{1}{k+1} \Big[z_0 g(\bar{i}_0, \bar{i}_1) + z_1 g(\bar{i}_1, \bar{i}_2) + \cdots z_{k-1} g(\bar{i}_{k-1}, \bar{i}_k) + z_k g(\bar{i}_k, \bar{i}_{k+1}) \Big]$$

$$d_{k-1}^{(\lambda)} = \frac{1}{k} \Big[z_0 g(\bar{i}_0, \bar{i}_1) + z_1 g(\bar{i}_1, \bar{i}_2) + \cdots z_{k-1} g(\bar{i}_{k-1}, \bar{i}_k) \Big]$$

$\therefore (1-\delta_k) d_{k-1}^{(\lambda)} + \delta_k z_k g(\bar{i}_k, \bar{i}_{k+1})$

$$= \frac{k}{k+1} d_{k-1}^{(\lambda)} + \frac{1}{k+1} z_k g(\bar{i}_k, \bar{i}_{k+1})$$

$$= \frac{\cancel{k}}{k+1} \times \frac{1}{\cancel{k}} \Big[z_0 g(\bar{i}_0, \bar{i}_1) + \cdots z_{k-1} g(\bar{i}_{k-1}, \bar{i}_k) \Big] + \frac{1}{k+1} z_k g(\bar{i}_k, \bar{i}_{k+1})$$

$$= \frac{1}{k+1} \Big[z_0 g(\bar{i}_0, \bar{i}_1) + \cdots z_k g(\bar{i}_k, \bar{i}_{k+1}) \Big] = d_k^{(\lambda)} \,/\!/$$

Prob4). $T^{(t+1)}J$를 $(t+1)$ stages와 terminal cost J를 갖는 MDP의 cost vector로 생각하면,

$$T^{t+1}J = \alpha^{t+1}P^{t+1}J + \sum_{k=0}^{t}\alpha^k P^k g$$

이고

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

이므로

$$\begin{aligned}(T^{(\lambda)}J)(i) &= \sum_{t=0}^{\infty}(1-\lambda)\lambda^t E\{\alpha^{t+1}J(i_{t+1}) + \sum_{k=0}^{t}\alpha^k g(i_k, i_{k+1}) \mid i_0 = i\} \\ &= J(i) + (1-\lambda)\sum_{t=0}^{\infty}\sum_{k=0}^{t}\lambda^t\alpha^k E\{g(i_k, i_{k+1}) + \alpha J_t(i_{k+1}) - J_t(i_k) \mid i_0 = i\} \\ &= J(i) + (1-\lambda)\sum_{k=0}^{\infty}\left(\sum_{t=k}^{\infty}\lambda^t\right)\alpha^k E\{g(i_k, i_{k+1}) + \alpha J(i_{k+1}) - J(i_k) \mid i_0 = i\} \\ &= J(i) + \sum_{t=0}^{\infty}(\alpha\lambda)^k E\{g(i_t, i_{t+1}) + \alpha J(i_{t+1}) - J(i_t) \mid i_0 = i\}\end{aligned}$$

PVI(λ)는 $\Phi r_{k+1} = \Pi T^{(\lambda)}(\Phi r_k)$ 이므로

$$r_{k+1} = \arg\min_{r\in R^s}\sum_{i=1}^{n}\xi_i\Big(\phi(i)'r - \phi(i)'r_k - \sum_{t=0}^{\infty}(\alpha\lambda)^t E\{\underbrace{g(i_t, i_{t+1}) + \alpha\phi(i_{t+1})'r_k - \phi(i_t)'r_k}_{d_k(i_t, i_{t+1})} \mid i_0 = i\}\Big)^2$$

$$= \arg\min_{r\in R^s}\sum_{i=1}^{n}\xi_i\Big(\phi(i)'r - \phi(i)'r_k - \sum_{t=0}^{\infty}(\alpha\lambda)^t E\{d_k(i_t, i_{t+1}) \mid i_0 = i\}\Big)^2 \quad (1)$$

LSPE(λ)는 PVI(λ)의 simulation-based approximation 이므로, 무한한 trajectory $(i_0, i_1, \dots)$을 만들고 (i_k, i_{k+1})로 transition이 일어날 때 r_k를 update해야한다

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

(1) 식을 미분하여 0으로 만드는 r을 찾아 대입하면

$$r_{k+1} = \left(\sum_{i=1}^{n}\xi_i\phi(i)\phi(i)'\right)^{-1}\left(\sum_{i=1}^{n}\xi_i\phi(i)\Big(\phi(i)'r_k + \sum_{t=0}^{\infty}(\alpha\lambda)^t E\{d_k(i_t, i_{t+1}) \mid i_0 = i\}\Big)\right)$$

(2)식에 대해서도 같은 방법으로하면,

$$\begin{aligned}r_{k+1} &= \left(\sum_{k=0}^{k}\phi(i_t)\phi'(i_t)\right)^{-1}\left(\sum_{t=0}^{k}\phi(i_t)\Big(\phi(i_t)r_k + \sum_{m=t}^{k}(\alpha\lambda)^{m-t}d_k(i_m, i_{m+1})\Big)\right) \\ &= \left(\sum_{i=1}^{n}\hat{\xi}_{i,k}\phi(i)\phi(i)'\right)^{-1}\left(\sum_{i=1}^{n}\hat{\xi}_{i,k}\phi(i)\Big(\phi(i)r_k + \sum_{m=t}^{k}(\alpha\lambda)^{m-t}d_k(i_m, i_{m+1})\Big)\right)\end{aligned}$$

$$\hat{\xi}_{i,k} = \frac{\sum_{t=0}^{k}\delta(i_t = i)}{k+1}$$

$k\to\infty$ 이면 $\hat{\xi}_{i,k}\to\xi_i$ 이고 (2) 식의 finite discounted sum $\to$ infinite discounted sum 이므로 (2) 식은 (1)의 approximation이다