Problem 1 (a)

proof : $V_*(s) = \max_a Q_*(s,a)$

$$V^{\pi}(s) = \sum_{a\in A}\left[\pi(a|s)\, Q^{\pi}(s,a)\right]$$

$$\leq \sum_{a\in A}\left[\pi(a|s) \max_a Q^{\pi}(s,a)\right]$$

$$= \max_a Q^{\pi}(s,a) \times \sum_{a\in A}\pi(a|s)$$

$$= \max_a Q^{\pi}(s,a)$$

$$\Rightarrow V^{\pi}(s) \leq \max_a Q^{\pi}(s,a) \quad \text{—— ①}$$

if $V^{\pi^*}(s) < \max_a Q^{\pi^*}(s,a)$

則必存在一 policy $\pi_{new} > \pi^*$ (與 $\pi^* \geq \pi$ 矛盾)

$$\pi_{new} = \begin{cases} \pi_{new}(s) = \arg\max_a Q^{\pi^*}(s,a), & \forall s \in S \\ \pi_{new}(\bar{s}) = \pi^*(\bar{s}), & \bar{s} \in S - s \end{cases}$$

$$\Rightarrow V^{\pi^*}(s) \geq \max_a Q^{\pi^*}(s,a) \quad \text{—— ②}$$

由 ① ② $\Rightarrow V^{\pi^*}(s) = \max_a Q^{\pi^*}(s,a)$ ※

Problem 1 (a)

proof : $Q_*(s,a) = R_s^a + \gamma \sum_{s'} P_{ss'}^a V_*(s')$

$\sum_{s'} P_{ss'}^a V_*(s') - \sum_{s'} P_{ss'}^a V^\pi(s)$

$= \sum_{s'} P_{ss'}^a [V_*(s') - V^\pi(s)] \geq 0$

$(\because \pi^* \geq \pi \ \therefore V_*(s') \geq V^\pi(s'))$

$\Rightarrow \sum_{s'} P_{ss'}^a V_*(s') \geq \sum_{s'} P_{ss'}^a V^\pi(s)$, for all π

$\Rightarrow \sum_{s'} P_{ss'}^a V_*(s') = \max_\pi \sum_{s'} P_{ss'}^a V^\pi(s)$ —— ①

$Q_*(s,a) = R_s^a + \max_\pi \gamma \sum_{s'} P_{ss'}^a V^\pi(s')$

$= R_s^a + \gamma \max_\pi \sum_{s'} P_{ss'}^a V^\pi(s')$

$= R_s^a + \gamma \sum_{s'} P_{ss'}^a V_*(s')$ …… 由 ①

※

Problem 1 (b)

$$\| T^*(Q) - T^*(Q') \|$$

$$= \max_{s,a} \left| [T^*(Q)](s,a) - [T^*(Q')](s,a) \right|$$

$$= \max_{s,a} \left| \left(\cancel{R_s^a} + \gamma \sum_{s'} P_{ss'}^a \max_{a'} Q(s',a') \right) - \left(\cancel{R_s^a} + \gamma \sum_{s'} P_{ss'}^a \max_{a'} Q'(s',a') \right) \right|$$

$$= \gamma \max_{s,a} \left| \sum_{s'} P_{ss'}^a \left[\max_{a'} Q(s',a') - \max_{a'} Q'(s',a') \right] \right|$$

$$\leq \gamma \max_{s,a} \left| \sum_{s'} P_{ss'}^a \max_{a'} \left[Q(s',a') - Q'(s',a') \right] \right|$$

$$\leq \gamma \max_{s,a} \left| \max_{s',a'} \left[Q(s',a') - Q'(s',a') \right] \right| \qquad \left(\because \sum_{s'} P_{ss'}^a = 1 \right)$$

$$= \gamma \max_{s,a} \left| Q(s,a) - Q'(s,a) \right|$$

$$= \gamma \| Q - Q' \|_\infty$$

※

Problem 2 (a)

(i)

$d_p(aU, aV)$

$$= \inf_{aU \sim F_{aU},\, aV \sim F_{aV}} \|aU - aV\|_p$$

$$= \inf_{aU \sim F_{aU},\, aV \sim F_{aV}} \|a(U-V)\|_p$$

$$= \inf_{aU \sim F_{aU},\, aV \sim F_{aV}} |a| \cdot \|U-V\|_p \quad \cdots \text{ absolute homogeneity}$$

$$= |a| \inf_{aU \sim F_{aU},\, aV \sim F_{aV}} \|U-V\|_p$$

$$= |a|\, d_p(U, V) \quad ※$$

Problem 2 (a)

(ii)

$d_P(A+U, A+V)$

$$= \inf_{A+U\sim F_{A+U},\, A+V\sim F_{A+V}} \|(A+U)-(A+V)\|_P$$

$$= \inf_{A+U\sim F_{A+U},\, A+V\sim F_{A+V}} |P(A+U) - P(A+V)|$$

$$= \inf_{A+U\sim F_{A+U},\, A+V\sim F_{A+V}} |P(A)P(U) - P(A)P(V)| \quad \because A \text{ is independent of } U, V$$

$$= \inf_{A+U\sim F_{A+U},\, A+V\sim F_{A+V}} |P(A)[P(U) - P(V)]|$$

$$\leq \inf_{U\sim F_U,\, V\sim F_V} |P(U) - P(V)| \quad \because P(A) \leq 1$$

$$= \inf_{U\sim F_U,\, V\sim F_V} \|U - V\|_P$$

$$= d_P(U, V) \quad ※$$

Problem 2 (a)

(iii)

$d_p(QU, QV)$

$$= \inf_{QU \sim F_{QU},\, QV \sim F_{QV}} \| QU - QV \|_p$$

$$= \inf_{QU \sim F_{QU},\, QV \sim F_{QV}} \| Q(U - V) \|_p$$

$$\leq \inf_{QU \sim F_{QU},\, QV \sim F_{QV}} \|Q\| \, \|U - V\|_p \quad \cdots \text{ Cauchy-Schwarz Inequality}$$

$$= \inf_{QU \sim F_{QU},\, QV \sim F_{QV}} P(Q=1) \|U - V\|_p + P(Q=0) \times 0$$

$$= q \cdot d_p(U, V) \quad \#$$

Problem 2 (b)

$$d_p(B^{\pi} Z_1(s,a), B^{\pi} Z_2(s,a))$$

$$= d_p(r(s,a) + \gamma P^{\pi} Z_1(s,a), r(s,a) + \gamma P^{\pi} Z_2(s,a))$$

$$\leq d_p(\gamma P^{\pi} Z_1(s,a), \gamma P^{\pi} Z_2(s,a)) \quad \cdots \text{由 (ii)}$$

$$= \gamma\, d_p(P^{\pi} Z_1(s,a), P^{\pi} Z_2(s,a)) \quad \cdots \text{由 (i)}$$

$$= \gamma \sup_{\bar{s}, \bar{a}} d_p(Z_1(\bar{s}, \bar{a}), Z_2(\bar{s}, \bar{a}))$$

※