Problem1

Property1 $L_{\pi_{\theta_1}}(\pi_{\theta_1}) = \eta(\pi_{\theta_1})$

pf $L_{\pi_{\theta_1}}(\pi_{\theta_1}) = \eta(\pi_{\theta_1}) + \sum_{s\in S} d_\mu^{\pi_{\theta_1}}(s) \underline{\sum_{a\in A} \pi_{\theta_1}(a|s) A^{\pi_{\theta_1}}(s,a)}$

(by definition of $L_{\pi_{\theta_1}}(\pi_\theta)$ in (1))

$= \eta(\pi_{\theta_1}) + \left(\sum_{s\in S} d_\mu^{\pi_{\theta_1}}(s)\right) \cdot 0$

$= \eta(\pi_{\theta_1})$

Claim: ___ $= 0$

___ $= \sum_{a\in A} \pi_{\theta_1}(a|s) \underline{(Q^{\pi_1}(s,a) - V^{\pi_1}(s,a))}$ $\quad$ $\because A^{\pi_{\theta_1}}(s,a)$ by defition of A^{π_θ}

$= \underline{\sum_{a\in A} \pi_{\theta_1}(a|s) Q^{\pi_1}(s,a)} - \underbrace{\sum_{a\in A} \pi_{\theta_1}(a|s)}_{=1} V^{\pi_1}(s,a)$

$= \underline{V^{\pi_1}(s)} - V^{\pi_1}(s)$

(by bellman equation)

$= 0$ $\square$

Property2 $\nabla_\theta L_{\pi_{\theta_1}}(\pi_\theta)|_{\theta=\theta_1} = \nabla_\theta \eta(\pi_\theta)|_{\theta=\theta_1}$

pf By Performance Lemma and (1), we have

$\eta(\pi_\theta) = \eta(\pi_{\theta_1}) + \underline{\sum_s d_\mu^{\pi_\theta}(s) \sum_a \pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)}$ RHS1

$L_{\pi_{\theta_1}}(\pi_\theta) = \eta(\pi_{\theta_1}) + \underline{\sum_s d_\mu^{\pi_{\theta_1}}(s) \sum_a \pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)}$ RHS2

$\Rightarrow$ If $\frac{\partial RHS1}{\partial\theta}\Big|_{\theta=\theta_1} = \frac{\partial RHS2}{\partial\theta}\Big|_{\theta=\theta_1}$, then Property2 holds.

$\frac{\partial RHS1}{\partial\theta}\Big|_{\theta=\theta_1} = \underbrace{\left(\sum_s \nabla_\theta d_\mu^{\pi_\theta}(s) \sum_a \pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)\right)\Big|_{\theta=\theta_1}}_{0}$

$+ \left(\sum_s d_\mu^{\pi_\theta}(s) \sum_a \nabla_\theta \left[\pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)\right]\right)\Big|_{\theta=\theta_1}$

(by Chain rule) $\quad$ $\because \sum_a \pi_{\theta_1}(a|s) A^{\pi_{\theta_1}}(s,a) = 0$

$$\frac{\partial RHS2}{\partial \theta}\Big|_{\theta=\theta_1} = \Big(\sum_s \nabla d_\mu^{\pi_{\theta_1}}(s) \sum \pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)\Big)\Big|_{\theta=\theta_1}$$

$$+\Big(\sum_s d_\mu^{\pi_{\theta_1}}(s) \sum_a \nabla_\theta \big[\pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)\big]\Big)\Big|_{\theta=\theta_1}$$

$$\Rightarrow \frac{\partial RHS1}{\partial \theta}\Big|_{\theta=\theta_1} = \frac{\partial RHS2}{\partial \theta}\Big|_{\theta=\theta_2}$$

$\Rightarrow$ Property2 holds

Problem2

(a) We use two Lemma to solve.

Lemma1

Let $f: \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = X^T A X$,
where A: symmetric and $X = (X_1 \cdots X_n)^T$.
Then $\frac{\partial f}{\partial X} = 2AX$.

pf $y = f(x) = X^T A X = \sum_{i=1}^n \sum_{j=1}^n a_{ij} X_i X_j$

$$= \sum_{i=1}^n a_{ip} X_i X_p + \sum_{j=1}^n a_{pj} X_p X_j + \sum_{\substack{i=1 \\ i,j \neq p}}^n \sum_{j=1}^n a_{ij} X_i X_j$$

$$\frac{\partial y}{\partial X_p} = \sum_{i=1}^n a_{ip} X_i + \sum_{j=1}^n a_{pj} X_j$$

$$= 2\sum_{i=1}^n a_{pi} X_i \quad (\because A: \text{symmetric})$$

$$\Rightarrow \frac{\partial f}{\partial X} = \begin{pmatrix} 2\sum_{i=1}^n a_{1i} X_i \\ \vdots \\ 2\sum_{i=1}^n a_{ni} X_i \end{pmatrix} = 2\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$= 2AX \quad \square$$

Lemma2

$$\boxed{\frac{\partial A^T x}{\partial x} = A, \text{ where } A\text{: metrix}}$$

pf similar to above proof. □

Let $\frac{\partial \mathcal{L}(\theta,\lambda)}{\partial \theta} = 0.$

$$\Rightarrow -(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k}) + \frac{\lambda}{\cancel{2}}(\cancel{2} H_{\theta_k}(\theta-\theta_k)) = 0$$

by Lemma1, Lemma2.

$$\Rightarrow (\theta-\theta_k) = \frac{1}{\lambda} H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k})$$ 代回 (4)

$$\Rightarrow \min_\theta \mathcal{L}(\theta,\lambda) = -(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})^T \frac{1}{\lambda} H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k}) + \lambda\left(\frac{1}{2}\left(\frac{1}{\lambda} H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k})\right)^T H_{\theta_k}(\theta-\theta_k) - \delta\right)$$

$$= -(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})^T \frac{1}{\lambda} H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k}) + \cancel{\lambda}\left(\frac{1}{2}(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})^T \cancel{(H^{-1}_{\theta_k})^T} \frac{1}{\cancel{\lambda}} \cancel{H_{\theta_k}}(\theta-\theta_k)\right) - \lambda\delta$$

∵ H: symmetric

$$\overset{\min}{=} -(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})^T \frac{1}{\lambda} H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k}) + \frac{1}{2}(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}) \frac{1}{\lambda} H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k}) - \lambda\delta$$

$$= -\frac{1}{2\lambda}(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})^T H^{-1}_{\theta_k}(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}) - \lambda\delta$$

$$= \min_{\theta\in\mathbb{R}^d} \mathcal{L}(\theta,\lambda) \quad \square$$

(∵ it's LP transformation, strong duality holds)

② Let $\frac{dD(\lambda)}{d\lambda}=0$

$$\Rightarrow \frac{dD(\lambda)}{d\lambda} = \frac{1}{2\lambda^2}\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right)^T H_{\theta_k}^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right) - \delta$$

$$\Rightarrow \left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right)^T H_{\theta_k}^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right) = 2\lambda^2\delta$$

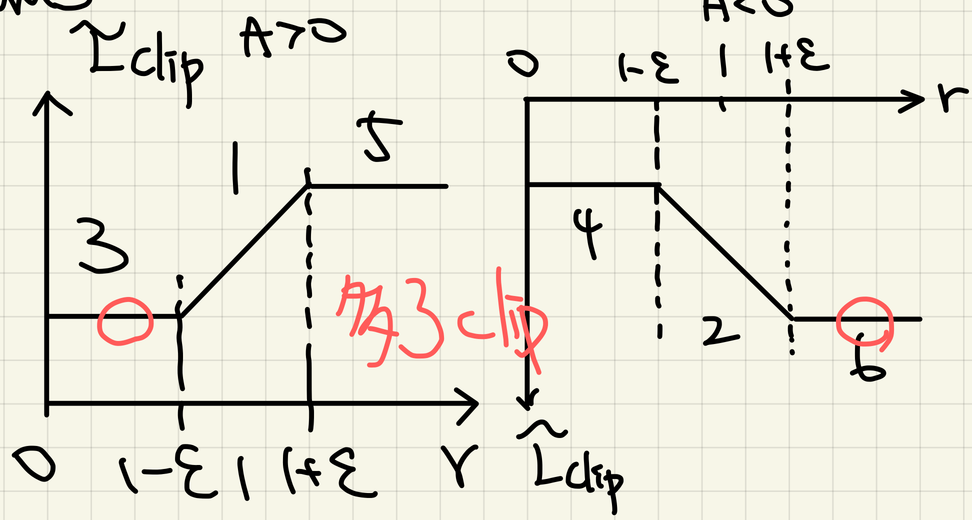
$$\Rightarrow \lambda^* = \sqrt{\frac{\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right)^T H_{\theta_k}^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right)}{2\delta}} \quad \square$$

(b) ① Let $\frac{\partial \mathcal{L}(\theta,\lambda)}{\partial \theta}=0$, we have

$$(\theta^*-\theta_k) = \frac{1}{\lambda^*} H_{\theta_k}^{-1}\left(\nabla_\theta L_{\theta_k}|_{\theta=\theta_k}\right)$$

② $$\alpha = \frac{1}{\lambda^*} = \sqrt{\frac{2\delta}{\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right)^T H_{\theta_k}^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}\right)}} \quad \square$$

Problem3



num	$p_t(\theta) > 0$	A_t	Return Value of min	Objective is Clipped	Sign of Objective	Gradient
1	$p_t(\theta) \in [1-\epsilon, 1+\epsilon]$	+	$p_t(\theta)A_t$	no	+	✅
2	$p_t(\theta) \in [1-\epsilon, 1+\epsilon]$	-	$p_t(\theta)A_t$	no	-	✅
3	$p_t(\theta) < 1-\epsilon$	+	$(1-\epsilon)p_t(\theta)A_t$	yes	+	0
4	$p_t(\theta) < 1-\epsilon$	-	$(1-\epsilon)p_t(\theta)A_t$	yes	-	0
5	$p_t(\theta) > 1+\epsilon$	+	$(1+\epsilon)p_t(\theta)A_t$	yes	+	0
5	$p_t(\theta) > 1+\epsilon$	-	$(1+\epsilon)p_t(\theta)A_t$	yes	-	0