Problem 1: $$\eta(\pi_\theta) := \eta(\pi_{\theta_1}) + \sum_s d_\mu^{\pi_\theta}(s) \cdot \sum_a \pi_\theta(a|s) \cdot A^{\pi_{\theta_1}}(s,a) \quad \text{— } (*)$$

$$L_{\pi_{\theta_1}}(\pi_\theta) := \underbrace{\eta(\pi_{\theta_1})}_{\text{not a function of } \theta} + \sum_s d_\mu^{\pi_{\theta_1}}(s) \cdot \sum_a \pi_\theta(a|s) \cdot A^{\pi_{\theta_1}}(s,a) \quad \text{— } (**)$$

$\boxed{\text{P. 1}}$

(i) $$L_{\pi_{\theta_1}}(\pi_{\theta_1}) = \eta(\pi_{\theta_1}) + \sum_s d_\mu^{\pi_{\theta_1}}(s) \cdot \underbrace{\sum_a \pi_{\theta_1}(a|s) \cdot A^{\pi_{\theta_1}}(s,a)}_{=0, \text{ for every } s}$$

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

(ii). By taking the gradients of both sides of $(*)$ and $(**)$, we have

$$\nabla_\theta L_{\pi_{\theta_1}}(\pi_\theta) = \sum_s d_\mu^{\pi_{\theta_1}}(s) \cdot \sum_a \nabla_\theta \pi_\theta(a|s) \cdot A^{\pi_{\theta_1}}(s,a).$$

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

Then, we know

$$\nabla_\theta \eta(\pi_\theta)\Big|_{\theta=\theta_1} = \sum_s \left(\nabla_\theta \left(d_\mu^{\pi_\theta}(s) \right)\Big|_{\theta=\theta_1} \cdot \underbrace{\left(\sum_a \pi_{\theta_1}(a|s) \cdot A^{\pi_{\theta_1}}(s,a) \right)}_{=0, \text{ for every } s} + d_\mu^{\pi_{\theta_1}}(s) \cdot \sum_a \nabla_\theta \pi_\theta(a|s)\Big|_{\theta=\theta_1} A^{\pi_{\theta_1}}(s,a) \right)$$

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

$$= \nabla_\theta L_{\pi_{\theta_1}}(\pi_\theta)\Big|_{\theta=\theta_1}.$$

$\square$

Problem 2

(a) Note that $\mathcal{L}(\theta,\lambda) := -\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T(\theta-\theta_k) + \lambda\cdot\left(\frac{1}{2}(\theta-\theta_k)^T H(\theta-\theta_k) - \delta\right)$

Since H is positive definite, we know $\mathcal{L}(\theta,\lambda)$ is strictly convex in θ.

To find $D(\lambda) = \min_\theta \mathcal{L}(\theta,\lambda)$, we simply take the gradient of $\mathcal{L}(\theta,\lambda)$ w.r.t. θ and set it to zero:

$$\nabla_\theta \mathcal{L}(\theta,\lambda) = \underbrace{-\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k} + \lambda\cdot H(\theta-\theta_k) = 0.}_{} \qquad \text{———} (*)$$

since H^{-1} exists, this equation has a unique solution.

Define $\bar{\theta}(\lambda) := \theta_k + \frac{1}{\lambda} H^{-1} \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}$ ——— $(**)$

This implies that
$$D(\lambda) = \mathcal{L}(\bar{\theta}(\lambda),\lambda)$$
$$= -\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T \cdot \frac{1}{\lambda} H^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)$$
$$+ \frac{1}{2}\frac{1}{\lambda}\cdot\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T \cdot H^{-1}\cdot\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right).$$
$$-\lambda\delta$$

Hence, we have $D(\lambda) = -\frac{1}{2\lambda}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T H^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right) - \lambda\delta.$ □

Moreover, by the Arithmetic-Geometric Mean Inequality, we know

$$-D(\lambda) = \frac{1}{2\lambda}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T H^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right) + \lambda\delta \geq 2\sqrt{\frac{\delta\cdot\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T H^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)}{2}},$$

and the equality holds when $\frac{1}{2\lambda}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T H^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right) = \lambda\delta.$

Hence, we know $\lambda^* := \arg\max_{\lambda\geq 0} D(\lambda) = \sqrt{\frac{\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)^T H^{-1}\left(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}\right)}{2\delta}}$ □

(b). By (a), we have $\lambda^* = \sqrt{\dfrac{(\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})^T H^{-1} (\nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k})}{2\delta}}$.

By combining the above and (*)-(**), we have

$$\theta^* := \arg\min_\theta \mathcal{L}(\theta, \lambda^*) = \theta_k + \frac{1}{\lambda^*} H^{-1} \nabla_\theta L_{\theta_k}(\theta)|_{\theta=\theta_k}$$

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

This suggests that the step size

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

$\square$