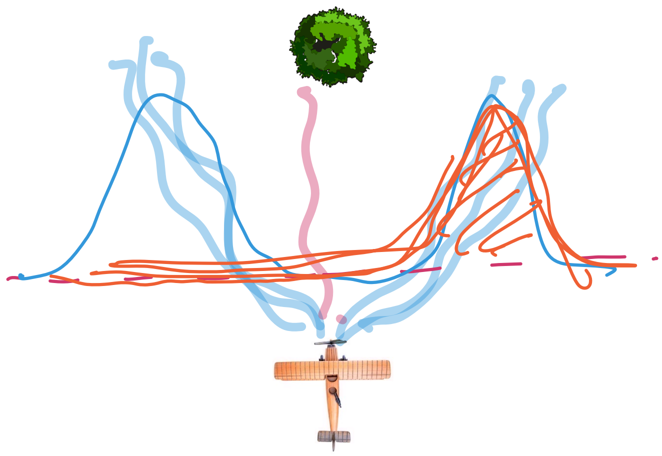
FORWARD KL

$$\min_{\theta} \; - \mathop{E}_{\xi \sim p_{expert}(\xi)} \log P_\theta(\xi)$$

$$\downarrow$$

$$- \sum_{\xi} p_{expert}(\xi) \log P_\theta(\xi)$$

$\neq 0$ $\qquad$ $= 0$



GOAL: LEARN $P_\theta(\xi)$

GIVEN:
- $\xi^h \sim P_{exper}(\xi^h)$
- BASIS FEATURE VECTORS

$f^1(\xi) \to \mathbb{R}$

$f^2(\xi) \to \mathbb{R}$

$\vdots$

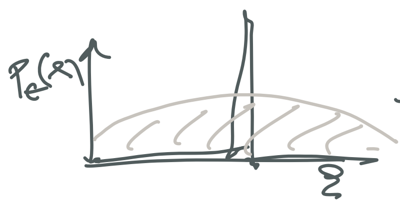
$f^k(\xi) \to \mathbb{R}$

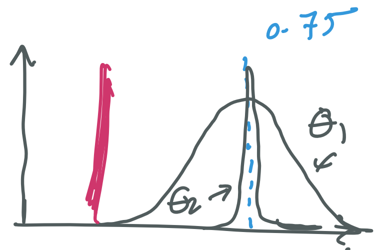
FIND $\max \; -\sum_{\xi} P_\theta(\xi) \log P_\theta(\xi)$ 0.75

s.t. $\sum_{\xi} P_\theta(\xi) f^1(\xi) = \mathop{E}_{\xi^h \sim P_{exp}} f^1(\xi^h)$

$\vdots$

$\sum_{\xi} P_\theta(\xi) f^k(\,) = E\, f^k(\,)$





$f^1(\xi)$ = How much grass ξ drives on?

$\min_\theta \sum_{\xi} P_\theta(\xi) \log P_\theta(\xi)$

s.t. $\sum_{\theta} P_\theta(\xi) f^1(\xi) = c_1 = 0.75$
0.55

① Write out the Lagrangian

$\max_\lambda \min_\theta \sum_\xi P_\theta(\xi) \log P_\theta(\xi)$

$+ \lambda_1 \left[\sum_\theta P_\theta f^1(\xi) - c \right]$

$+ \lambda_2 [\,- - -\,]$

$+ \lambda_3 [\,- - -\,]$

② Take the gradient

$\frac{\partial}{\partial \lambda}(\cdot) = 0$

③ Substitute in

$\boxed{\sum_\xi P_\theta(\xi) = 1}$

$$P_\theta(\xi) \propto \exp\left(-\lambda_1 f_1(\xi) - \lambda_2 f_2(\xi) - \lambda_3 f_3(\xi) \cdots\right)$$

$$\propto \exp\left(-\mathrm{cost}_\theta(\xi)\right)$$

$$= \frac{1}{Z(\theta)} \exp\left(-\mathrm{cost}_\theta(\xi)\right)$$

Plug back in to optimization

$$\max_\theta \sum_{i=1}^{N} \log P_\theta(\xi_i^h)$$

$$\sum_{i=1}^{N} \log\left(\frac{1}{Z(\theta)} \exp\left(-\mathrm{cost}_\theta(\xi_i^h)\right)\right)$$

$$= \sum_{i=1}^{N} -\mathrm{cost}_\theta(\xi_i^h) - \log Z(\theta)$$

$$\min_\theta \sum_{i=1}^{N} \mathrm{cost}_\theta(\xi_i^h) + \log Z(\theta)$$