$$\max_{P(\xi)} \; -\sum_{\xi} P(\xi) \log P(\xi)$$

$$\text{s.t.} \quad \sum_{\xi} P(\xi) f^k(\xi) = \underbrace{\sum_{\xi} P_{exp}(\xi) f^k(\xi)}_{\text{Expert feats.}}$$

$$= F^k \qquad \forall k = 1, \ldots K$$

λ^k for each constraint.

$$\mathcal{L}(\cdot) = -\sum_{\xi} P(\xi) \log P(\xi) - \sum_{k} \lambda^k \left(\sum_{\xi} P(\xi) f^k(\xi) - F^k \right)$$

LAGRANGE MULTIPLIER — CONSTRAINT VIOLATION

TAKE GRADIENT W.R.T $P(\xi)$.

$$-\left(P(\xi) \cdot \frac{1}{P(\xi)} + 1 \cdot \log P(\xi) \right) - \sum_{k} \lambda^k \cdot \left(f^k(\xi) \right) = 0$$

$$-1 - \log P(\xi) - \sum_{k} \lambda^k f^k(\xi) = 0$$

$$\log P(\xi) = -\sum_{k} \lambda^k f^k(\xi) - 1$$

$$\boxed{P(\xi) = \exp\left(-\sum_{k} \lambda^k f^k(\xi) - 1 \right)}$$

$$\sum_{\xi} P(\xi) = 1$$

$$P(\xi) = \frac{\exp\left(-\sum_k \lambda^k f^k(\xi)\right)}{Z(\lambda^k)}$$

$-$ COST OF A TRAJECTORY

$$Z(\lambda^k) = \sum_\xi \exp\left(-\sum_k \lambda^k f^k(\xi)\right)$$

BOLTZMANN DISTRIBUTION

$$P(\xi) = \frac{1}{Z}\exp\left(-C(\xi)\right)$$

$$\max_{P(\xi)} \; \mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \log P(\xi^h)$$

$$\max_{C()} \; \mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \log\left(\frac{1}{Z}\exp\left(-C(\xi^h)\right)\right)$$

$$\max_{\theta} \; \mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \log\left(\frac{1}{Z(\theta)}\exp\left(-C_\theta(\xi^h)\right)\right)$$

$$\nabla_\theta \mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \left[\log\left(\exp\left(-C_\theta(\xi^h)\right)\right) - \log Z(\theta)\right]$$

$$\mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \left[-\nabla C_\theta(\xi^h)\right] - \frac{1}{Z}\int \exp\left(-C_\theta(\xi)\right)\nabla_\theta C_\theta(\xi)$$

$$\mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \nabla_\theta C_\theta(\xi^h) \quad \frac{1}{Z(\theta)} \sum_\xi \nabla_\theta C_\theta(\xi)$$

$$P_\theta(\xi) = \frac{1}{Z(\theta)} \exp\left(-C_\theta(\xi)\right) \qquad Z(\theta) = \sum_\xi \exp(-C_\theta(\xi))$$

$$\nabla_\theta Z(\theta) = -\sum_\xi \exp(-C_\theta(\xi))\, \nabla_\theta C_\theta(\xi)$$

$$\mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} - \nabla_\theta C_\theta(\xi^h) + \frac{1}{Z(\theta)} \sum_\xi \exp(-C_\theta(\xi))\, \nabla_\theta C_\theta(\xi)$$

$$+ \sum_\xi P_\theta(\xi)\, \nabla_\theta C_\theta(\xi)$$

$$+ \mathop{E}_{\xi \sim P_\theta(\xi)} \nabla_\theta C_\theta(\xi)$$

$$\text{Gradient} = \mathop{E}_{\xi \sim P_\theta(\xi)} \nabla_\theta C_\theta(\xi) - \mathop{E}_{\xi^h \sim P_{exp}(\xi^h)} \nabla_\theta C_\theta(\xi^h)$$

↑↑↑ ↓↓↓


COST ON LEARNER
COST ON EXPERT TRAJ