Continuing on natural gradients.

Use distance measures other than Euclidean distances.

\[ C \rightarrow \sqrt{\Delta^T G \Delta} \rightarrow \sqrt{\Delta^T G(\theta) \Delta} \]

- adjusts the metric
- can vary according to \( \theta \)
- invertible + all eigenvalues positive

Note: \( G \) positive definite \( \Rightarrow \forall x: x^T G x > 0 \)
(positive semi-definite also works, but messier to work out)

Want: \( \arg \max \Delta : \Delta^T G(\theta) \Delta = 1 \)

\[ \Delta^T \nabla f(\theta) \]

\[ L(\Delta, \lambda) = \Delta^T \frac{\partial f(\theta)}{\partial \theta} - \lambda (\Delta^T G(\theta) \Delta) \]

\[ 0 = \frac{\partial L(\Delta, \lambda)}{\partial \Delta} = \frac{\partial f(\theta)}{\partial \theta} - 2 \lambda \Delta^T G(\theta) \]

\[ \Delta = G^{-1}(\theta) \frac{\partial f(\theta)}{\partial \theta} 2 \lambda \]-ignore, since a constant \( \lambda \)

\[ \hat{\nabla} f(\theta) = G^{-1}(\theta) \nabla f(\theta) \]

natural gradient

\[ \nabla f(\theta) \]

Euclidean gradient

Angle between these guaranteed to be < 90°
What if $f$ is a loss function that depends on a parameterized function probability distribution $d_\theta$?

Gradient descent on $d_\theta$, not on $\Theta$!

E.g. $L(d_\theta) = N(y, \sigma^2)$

Idea: Use Kullback–Leibler divergence (K-LD) as our "squared distance"

$$D_{KL}(q || p) = \sum_x q(x) \log \frac{q(x)}{p(x)}$$

Not really a distance! (Not symmetric, not the only choice.)

Q: How can we recover the $\Theta$ parameters when we move in function space?

In general, we can't.

$$\nabla f(\Theta) = \lim_{\varepsilon \to 0} \arg \max_{\Delta : D_KL(d_\theta \| d_\theta + \Delta) = 1} f(\Theta + \varepsilon \Delta)$$

Apply Taylor series expansion

$$g_q(\Theta) = D_{KL}(q || d_\theta$$

$$g_{d_\theta}(\Theta + \Delta) = D_{KL}(d_\theta \| d_{\Theta + \Delta})$$

$$\frac{\partial g_q(\Theta)}{\partial \Theta} = \frac{\partial}{\partial \Theta} \left( \sum_x q(x) \log \frac{q(x)}{d_\theta(x)} \right)$$

$$= \sum_x q(x) \frac{1}{q(x)} \frac{\partial q(x)}{\partial \Theta} - \sum_x \frac{q(x)}{d_\theta(x)} \frac{\partial d_\theta(x)}{\partial \Theta}$$

$$= \sum_x \frac{q(x)}{\partial \Theta} \left[ -\frac{q(x)}{d_\theta(x)} \frac{\partial d_\theta(x)}{\partial \Theta} + \frac{d_\theta(x)}{d_\theta(x)^2} \frac{\partial^2 d_\theta(x)}{\partial \Theta} \right]$$

$$= \sum_x \frac{q(x)}{\partial \Theta} \frac{\partial d_\theta(x)}{\partial \Theta}$$
\[ \frac{d^2 g_q(x)}{d \theta^2} = \sum_x -q(x) \frac{d^2 \theta(x)}{d \theta^2} + \frac{d d \theta(x)}{d \theta} \left( \frac{d \ln d \theta(x)}{d \theta} \right)^T \]

\[ = \sum_x -q(x) \frac{d^2 \theta(x)}{d \theta^2} + q(x) \left( \frac{d \ln d \theta(x)}{d \theta} \right)^T \left( \frac{d \ln d \theta(x)}{d \theta} \right) \]

\[ g_q(\theta, \Delta) = g_q(\theta) + \Delta^T \frac{d g_q(\theta)}{d \theta} + \frac{1}{2} \Delta^T \frac{d^2 g_q(\theta)}{d \theta^2} \Delta \]

Evaluate at \( \theta_o \):

\[ g_{d \theta}(\theta_o) = g_{d \theta}(\theta) + \Delta^T \left( \sum_x -q(x) \frac{d d \theta(x)}{d \theta} \right) + \frac{1}{2} \Delta^T \left( \sum_x -q(x) \frac{d^2 d \theta(x)}{d \theta^2} + \sum_x \frac{q(x)}{d \theta} \left( \frac{d \ln d \theta(x)}{d \theta} \right)^T \right) \Delta \]

\[ g_{d \theta}(\theta) = D_{KL}(d \theta \| d \theta_o) = 0 \]

\[ \sum_x \frac{d d \theta(x)}{d \theta} = \frac{d}{d \theta} \sum_x \frac{d \theta(x)}{d \theta} = 2 \theta_o \]

\[ \frac{1}{2} \Delta^T \left( \sum_x -q(x) \frac{d d \theta(x)}{d \theta} \right) \Delta \]

\[ \rightarrow \text{Natural gradient using FIM is covariant (also called invariant, or invariant to reparameterization).} \]

\[ \text{If two parameterizations can represent the same distributions, then using NF and small enough step size, one will get the same sequence of distributions.} \]

**Fisher Information Matrix (FIM)**: because we sum over a probability distribution.
Natural vs. Newton's method?

- Use Newton if you know the Hessian of the loss function $L$.
- Use NJ if you don't know the Hessian (or maybe not even $L$) but do know $J$.

Example: In RL: $J(s,a)$ → don't know the Hessian.

Can think of $\pi$ as a distribution over state trajectories — or as a collection of per-state distributions over actions.

Kakade 2002

$F(\theta) = \sum_{s} \pi(s) \sum_{a} \pi(a|s,\theta) \left( \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta} \cdot \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta} \right)^T$

estimate this know these

$\nabla J(\theta) = \frac{\partial}{\partial \theta} \sum_{s} \ln \pi(s,a,\theta) \cdot \frac{d \ln \pi(s,a,\theta)}{d \theta} \cdot \frac{d \ln \pi(s,a,\theta)}{d \theta}^T$

$\nabla J(\theta) = W$

Natural Actor-Critic

1) Approx $q \approx \nabla w(s)$ by $W \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta} + \nabla w(s)$

2) $\theta \leftarrow \theta + \Delta \theta$