687 2017-11-30

Note Title 11/30/2017

Continuing on natural gradients.

Use distance measures other than Euclidean distances.

$\hookrightarrow \sqrt{\Delta^T G \Delta} \rightsquigarrow \sqrt{\Delta^T G(\theta) \Delta}$

adjusts the metric

can vary according to θ

Note: G positive definite $\rightarrow \forall x: x^T G x > 0$

(positive semidefinite also works, but messier to work out) $\rightarrow \geq$

invertible + all eigenvalues positive

Want: $\underset{\Delta:\ \Delta^T G(\theta)\Delta = 1}{\arg\max} \ \Delta^T \frac{\partial f(\theta)}{\partial \theta}$

$$L(\Delta, \lambda) = \Delta^T \frac{\partial f(\theta)}{\partial \theta} - \lambda\left(\Delta^T G(\theta) \Delta\right)$$

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

$$\Delta = G^{-1}(\theta) \frac{\partial f(\theta)}{\partial \theta} \frac{1}{2\lambda}$$

$\leftarrow$ ignore, since a constant

See paper on GeNGA.

$$\tilde{\nabla} f(\theta) = G^{-1}(\theta) \nabla f(\theta)$$

natural gradient

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_θ?

Gradient descent on d_θ, not on θ!

E.g. $L(d_\theta)$ $\quad d_\theta = N\underbrace{(\mu, \sigma^2)}_{\theta}$

Idea: Use Kullback-Liebler 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 θ parameters when we move in function space?

In general we can't.

$$\nabla f(\theta) = \lim_{\epsilon \to 0} \operatorname*{argmax}_{\Delta : D_{KL}(d_\theta \| d_{\theta+\Delta}) = 1} f(\theta + \epsilon\Delta)$$

Apply Taylor series expansion

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

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

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

$$= \sum_x \cancel{q(x)} \frac{d_\theta(x)}{\cancel{q(x)}} \frac{\partial}{\partial \theta} \frac{q(x)}{d_\theta(x)}$$

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

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

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

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

$$g_q(\theta+\Delta) \overset{\text{Taylor}_2}{=} g_q(\theta) + \Delta^T \frac{\partial g_q(\theta)}{\partial\theta} + \frac{1}{2}\Delta^T \frac{\partial^2 g_q(\theta)}{\partial\theta} \Delta$$

$$= g_q(\theta) + \Delta^T\left(\sum_x \frac{-q(x)}{g_q(x)} \frac{\partial d_\theta(x)}{\partial\theta}\right) + \frac{1}{2}\Delta^T\left(\sum_x \frac{-q(x)}{d_\theta(x)} \frac{\partial^2 d_\theta(x)}{\partial\theta^2} + \sum_x q(x) \left(\frac{\partial \ln d_\theta(x)}{\partial\theta}\right)\left(\frac{\partial \ln d_\theta(x)}{\partial\theta}\right)^T\right)\Delta$$

Evaluate $\theta+\Delta$:

$$g_{d_\theta}(\theta+\Delta) = g_{d_\theta}(\theta) + \Delta^T\Bigg(\underbrace{\sum_x \frac{-d_\theta(x)}{d_\theta(x)} \frac{\partial d_\theta(x)}{\partial\theta}}_{\downarrow\, 0}\Bigg) + \frac{1}{2}\Delta^T\Bigg(\underbrace{\sum_x \frac{-d_\theta(x)}{d_\theta(x)} \frac{\partial^2 d_\theta(x)}{\partial\theta^2}}_{\downarrow\, 0} + \overbrace{\sum_x d_\theta(x) \left(\frac{\partial \ln d_\theta(x)}{\partial\theta}\right)\left(\frac{\partial \ln d_\theta(x)}{\partial\theta}\right)^T}^{NG}\Bigg)\Delta$$

$$g_{d_\theta}(\theta) = D_{KL}(d_\theta \| d_\theta) = 0$$

$$\sum_x \frac{\partial d_\theta(x)}{\partial\theta} = \frac{\partial}{\partial\theta}\sum_x d_\theta(x) = \frac{\partial}{\partial\theta} 1^{*}$$

$$\underbrace{\frac{1}{2}\Delta^T \sum_x d_\theta(x) \left(\frac{\partial \ln d_\theta(x)}{d_\theta}\right)\left(\frac{\partial \ln d_\theta(x)}{\partial\theta}\right)^T \Delta}_{\uparrow}$$

Fisher Information Matrix (FIM)

*: because we sum over a probability distribution.

Natural gradient using FIM is covariant (also called invariant, or invariant to reparameterization).

If two parameterizations can represent the same distributions, then using NG and small enough step size, one will get the same sequence of distributions.

Natural vs. Newton's method?

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

$\downarrow$

Example: In RL: $J(\pi_\theta) \rightarrow$ don't know the Hessian.

Can think of π as a distribution over state trajectories - or as a collection of per-state distributions over actions.

Kakade 2002

$$F(\theta) = \sum_s \underbrace{d^\pi(s)}_{\text{estimate this}} \underbrace{\sum_a \pi(s,a,\theta)\left[\frac{\partial \ln \pi(s,a,\theta)}{\partial \theta} \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta}^T\right]}_{\text{know these}}$$

NG in RL is

$$\tilde{\nabla} J(\theta) = (\cdot)^{-1} \sum_s d^\pi(s) \sum_a \pi(s,a,\theta) \cdot \left(w^T \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta}\right) \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta}$$

$$= (\cdot)^{-1} \left(\sum_s - \cdot \sum_a - \cdot \frac{\partial \ln(\pi(s,a,\theta))}{\partial \theta} \frac{\partial \ln(\pi(s,a,\theta))}{\partial \theta}^T\right) w$$

$\downarrow$

cancel, cancel, cancel!

$\downarrow$

$\tilde{\nabla} J(\theta) = w$

Natural Actor-Critic

1) Approx $q^\pi(s,a)$ by $w^T \frac{\partial \ln \pi(s,a,\theta)}{\partial \theta} + \hat{V}_w(s)$
2) $\theta \leftarrow \theta + \alpha w$