2. Very Basic of Information Geometry for TRPO

# number	2
Select	appendix

Introduction

In Trust-Region Policy Optimization(TRPO), some concepts of information geometry are used to update the parameters efficiently. This post is written to introduce those algorithms briefly to **help** understand TRPO. If you are already familiar with

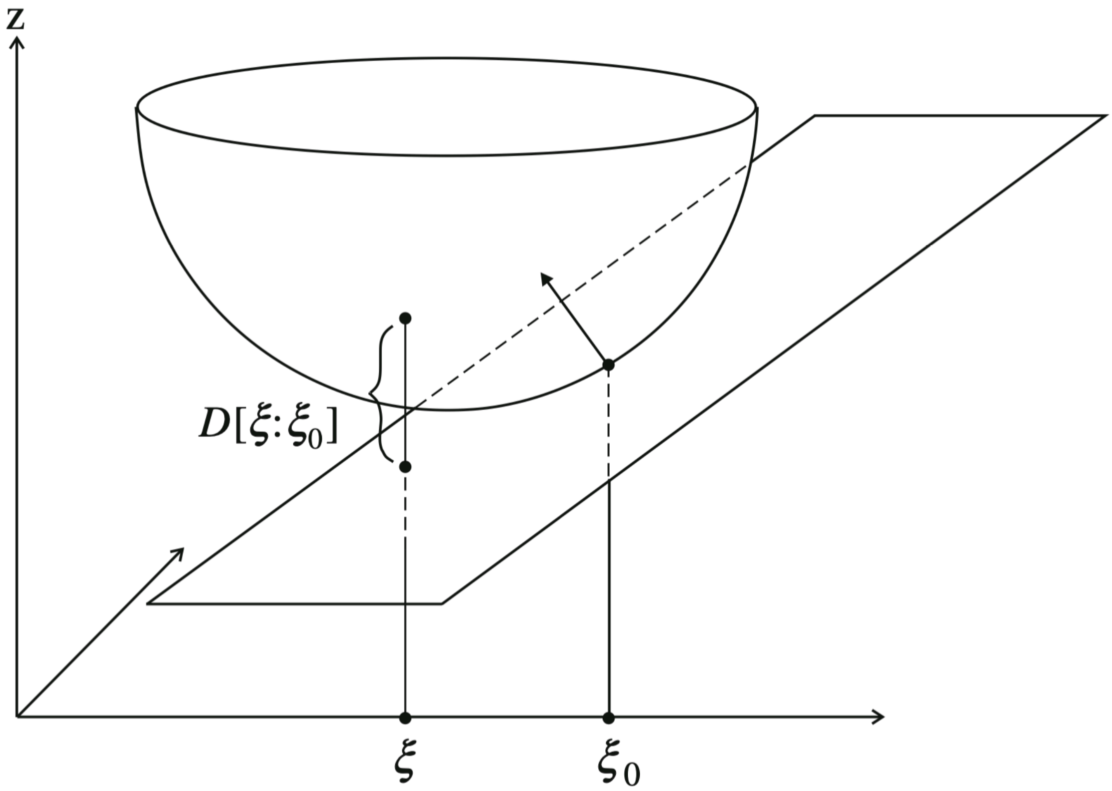
KL Divergence, Metric Tensor, Fisher Information Matrix, Natural (Policy) Gradient,

you may skip this post.

---

1. KL Divergence

Bregman Divergence and KL Divergence



Bregman Divergence

Consider a *convex function* ψ and its hyperplane touching it at ξ_0. We evaluate how much higher $\psi(\xi)$ relative to this tangent hyperplane. The difference is written as:

$$D_\psi[\xi : \xi_0] = \psi(\xi) - \psi(\xi_0) - \nabla\psi(\xi_0) \cdot (\xi - \xi_0)$$

This is called the *Bregman divergence* derived from a convex function ψ.

Now consider a specific convex function

$$\psi(\xi) = -\sum_{i=1}^{n} \xi_i \log \xi_i,$$

which is essentially an entropy-like function.

The corresponding Bregman divergence is

$$D_\psi[\xi : \xi'] = \sum_{i=1}^{n} \left(\xi_i \log \frac{\xi_i}{\xi'_i} - \xi_i + \xi'_i \right).$$

When $\sum \xi_i = \sum {\xi_i}' = 1$, it reduces to the *Kullback-Leibler Divergence (KL Divergence)*:

$$D_{KL}(p : q) = \sum_i p(i) \log \frac{p(i)}{q(i)}$$

where $p(i)$ and $q(i)$ are the discrete probability distributions.

Properties of KL Divergence

1. $D_{KL}(P:Q) \geq 0$, equality holds if and only if $P = Q$
2. In general, $D_{KL}(P:Q) \neq D_{KL}(Q:P)$
3. KL divergence is coordinate independent (invariant under reparameterization).
4. KL divergence is jointly convex.

From property 1 with uniform distribution Q, it can be shown that entropy of the arbitrary probability distribution P with dimension N is equal or less than $\log N$.

---

2. Metric Tensor

2.1 Definition of Metric Tensor

A metric tensor on an n**-dimensional manifold** is a symmetric $(0,2)$-tensor that defines the **inner product** (and hence "distances") between tangent vectors. Concretely, in local coordinates, it can be represented by an $n \times n$ **positive-definite (PSD) matrix** M. Given two vectors p and q in that coordinate system,

$$p \cdot q = p^T M q$$

For the basis vector e_i and e_j,

$$e_i \cdot e_j = g_{ij}$$

where g_{ij} is (i,j) component of the metric tensor M. In two-dimension Cartesian coordinates, metrix tensor is simply the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

2.2 Metric Tensor in Polar Coordinates

The transformation between Cartesian (x,y) and polar (r,θ) is given by:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

The transformation of the basis vectors is expressed with Jacobian:

$$\begin{pmatrix} e_r \\ e_\theta \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{pmatrix} \begin{pmatrix} e_x \\ e_y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} e_x \\ e_y \end{pmatrix}$$

With the transformation above, we can construct the metric tensor as $\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$

2.3 Distance and Volume

*Einstein Summation Convention

We will use Einstein summation convention. When the same index appears twice in one term, summation is automatically taken over this index even without the summation symbol.

$A = A^i e_i$ means $A = \sum A^i e_i$ unless specified otherwise.

Length and Volumes

Using Einstein summation convention, an infinitesimal line element ds is given as:

$$ds^2 = g_{ij} dx^i dx^j$$

Let g be the **determinant** of the metric tensor, then the volume element is:

$$dV = \sqrt{g} dx^1 dx^2 \cdots$$

For example, the length element in the polar coordinates is given as $ds^2 = dr^2 + r^2 d\theta^2$, and the volume(area) element is $dA = r dr d\theta$

---

3 Statistical Manifold and FIM as Metric Tensor

3.1 Statistical Manifold

A *statistical manifold* is a differentiable manifold whose each point corresponds to a probability distribution. For example, consider the family of univariate Gaussian distributions: $\mathcal{N}(\mu, \sigma^2)$, indexed by coordinates $\xi = (\mu, \sigma^2)$, $\sigma > 0$.

Rather than measuring the distance between two points (μ_1, σ_1) and (μ_2, σ_2) simply by Euclidean norm $\sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1^2 - \sigma_2^2)^2}$, one typically wants a measure that reflects how different the corresponding distributions are. This naturally motivates using *KL divergence* (or a symmetrized version).

3.2 KL Divergence and Fisher Information Metric

Line Element of Statistic Manifold

Let us define the manifold M whose point denotes probability distribution $p(x, \theta)$ with the parameter $\theta = (\theta^1, \theta^2, ..., \theta^n)$. The distance of two points on M is defined as KL divergence.

Strictly, KL divergence is a *'divergence'*, not a *'distance'* since it is not commutable:

$$D(p:q) \neq D(q:p).$$

Instead, we can define the *'distance' as* $D(p:q) = \frac{1}{2}\left(D_{KL}(p:q) + D_{KL}(q:p)\right)$. However, it gives same result as what we discuss in this chapter. So, we will use the term distance freely.

For convenience, we denote $p(x, \theta)$ by θ and D_{KL} as D.

The distance of the two *sufficiently close* point can be approximated by Taylor series:

$$D(\theta:\theta') \approx D(\theta':\theta') + \nabla_\theta D(\theta:\theta')^T|_{\theta=\theta'}(\theta-\theta') + \frac{1}{2}(\theta-\theta')^T \nabla_\theta^2 D(\theta:\theta')|_{\theta=\theta'}(\theta-\theta')$$

From simple calculation, it can be easily shown that first two terms equal to zero - that is:

$$D(\theta:\theta') = \frac{1}{2}(\theta-\theta')^T \nabla_\theta^2 D(\theta:\theta')|_{\theta'=\theta}(\theta-\theta') + O(\|\Delta\theta\|^3)$$

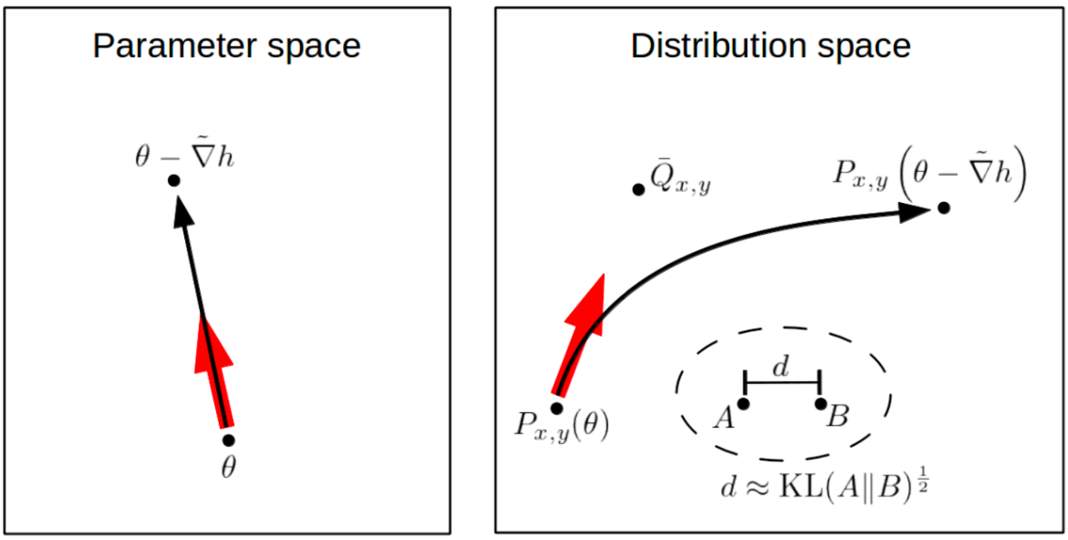
For a matrix $F = \nabla_\theta^2 D(\theta:\theta')|_{\theta'=\theta}$, the infinitesimal line element ds can be written as:

$$ds = D(\theta:\theta+d\theta) = \frac{1}{2} F_{ij} d\theta^i d\theta^j$$

To express in metric form, we can simply modify the distance to square root of KL divergence and multiply an appropriate scalar. Then, the length element is:

$$ds^2 = F_{ij} d\theta_i d\theta_j.$$

Now, we can define the matrix F as the metric tensor of manifold M.



Fisher Information Matrix

The matrix F defined above is called **Fisher Information Matrix(FIM)**.

FIM is defined in three ways and they both have the same results.

FIM is defined as:

1. Covariant of the Score Function
2. Expected Hessian of Negative Log-likelihood (NLL)
3. Hessian of KL Divergence

By the definition, FIM is calculated as:

$$F = \mathrm{E}_{x \sim p_\theta}[\partial_i \log p(x, \theta) \partial_j \log p(x, \theta)] = \mathrm{E}_{x \sim p_\theta}[-\nabla^2 \log p(x, \theta)]$$

FIM of an Univariate Gaussian Distribution

Let us calculate the FIM of an univariate Gaussian distribution:

$$f(x : \mu, v) = \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-\mu)^2}{2v}}.$$

NLL is:

$$L := -\log f = \frac{(x-\mu)^2}{2v} + \frac{\log v}{2} + \log \sqrt{2\pi}$$

To find Hessian of NLL:

$$\nabla^2_\mu L = \frac{1}{v}$$

$$\nabla^2_{\mu v} L = \nabla^2_{v\mu} L = -\frac{x-\mu}{v^2}$$

$$\nabla^2_v L = \frac{(x-\mu)^2}{v^3} - \frac{1}{2v^2}$$

By calculating the expectation, we can get FIM of f:

$$F = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}$$

If we construct a manifold M with this metric, the line element becomes

$$ds^2 = \frac{1}{\sigma^2} d\mu^2 + \frac{1}{2\sigma^4} d(\sigma^2)^2$$

or

$$ds^2 = \frac{1}{v}d\mu^2 + \frac{1}{2v^2}dv^2$$

The length of a curve $C = (\mu(\lambda), v(\lambda))$ on a manifold M is:

$$l = \int \sqrt{(ds^2)}$$

$$= \int \sqrt{\left(\frac{1}{v}d\mu\right)^2 + \left(\frac{1}{2v^2}dv\right)^2}$$

$$= \int_{\lambda_0}^{\lambda_1} \sqrt{\left(\frac{\mu'(\lambda)}{v(\lambda)}\right)^2 + \left(\frac{v'(\lambda)}{2\{v(\lambda)\}^2}\right)^2}\, d\lambda$$

4. Natural Gradient

4.1 Steepest Descent in Statistic Manifold

Given a function $L(\xi)$ in a manifold M with metric tensor F, let us change the current point ξ to $\xi + d\xi$, and see how the value of $L(\xi)$ changes, depending on the direction $d\xi$. We search for the direction in which L changes most rapidly with the step-size $\|d\xi\|^2 = g_{ij}d\xi^i d\xi^j$ being fixed to ϵ^2.

We put $d\xi = \epsilon a$ and require that $\|a\|^2 = g_{ij}a^i a^j = 1$.

Then, the steepest direction of L is the maximizer of

$$L(\xi + d\xi) - L(\xi) = \epsilon \nabla L(\xi) \cdot a$$

under the constraint of $\|a\| = 1$.

By Lagrange multiplier, this problem equivalents to maximizing $\epsilon \nabla L(\xi) \cdot a - \lambda g_{ij} a^i a^j$ with a
.

Therefore, the steepest direction is obtained as

$$a \propto F^{-1} \nabla L(\xi)$$

We call

$$\tilde{\nabla} L(\xi) = F^{-1}(\xi) \nabla L(\xi)$$

the *natural gradient* of L, where $\tilde{\nabla} = F^{-1}\nabla$ is the natural gradient operator.

In the univariate Gaussian manifold, the natural gradient operator is

$$\tilde{\nabla}_\mu = \sigma^2 \nabla_\mu$$

$$\tilde{\nabla}_{\sigma^2} = 2\sigma^4 \nabla_{\sigma^2}.$$

In general, computing F^{-1} is computational heavy especially when the dimension of F is large.

In this case, one can leverage conjugate gradient methods to calculate the inverse of F efficiently.

4.2 Natural Policy Gradient

Let us consider a system having state space $X = \{\mathbf{x}\}$ and action space $U = \{\mathbf{u}\}$. the FIM at the current state x is defined by:

$$\mathbf{F}(\theta|x) = \int \pi(u|x)\nabla_\theta \log \pi(u|x)\{\nabla_\theta \log \pi(u|x)\}^T du$$

where x and u are size n vectors.

The entire FIM is its expectation along all the trajectories,

$$\mathbf{G}(\theta) = \int d^\pi(x)\mathbf{F}(\theta|x)dx.$$

The natural gradient method, called the *natural policy gradient* or *natural actor-critic*, is given by

$$\theta_{t+1} = \theta_t + \eta G^{-1}(\theta_t)\nabla_\theta J(\theta_t).$$

It is known that natural gradient is saturation free: it is safe from the gradient vanishing problem.

Manifold Optimization

Finally, once a Riemannian structure (metric) is identified—be it via the Fisher Information Matrix or otherwise—one can generalize many standard Euclidean optimization algorithms (e.g., gradient descent, Newton's method) to *manifold optimization* methods, by replacing gradients and Hessians with their "natural" (i.e., covariant) versions on the manifold.