Riemannian Optimization

MLRG 28/07/21

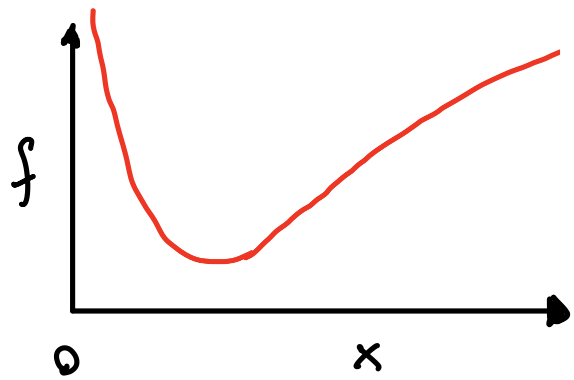
Riemannian Optimization

Riemannian Optimization

Riemannian Optimization

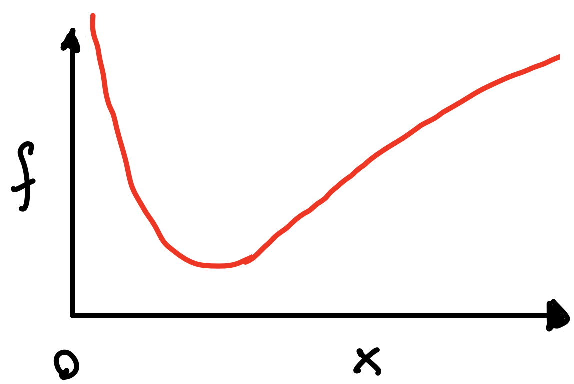
Some Problems

$$\min_{x > 0} f(x)$$

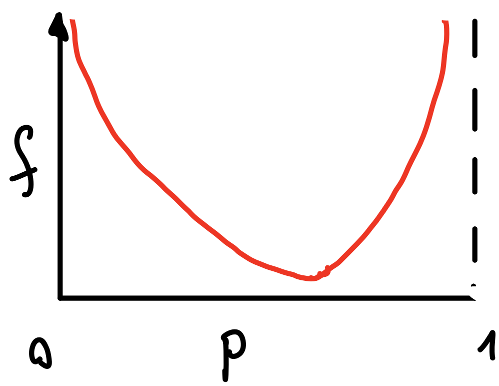


Some Problems

$$\min_{x>0} f(x)$$

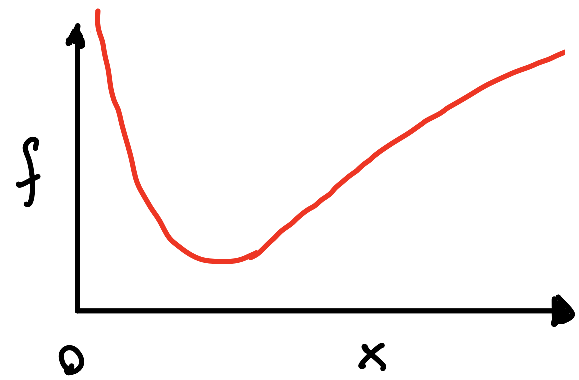


$$\min_{0<p<1} f(p)$$

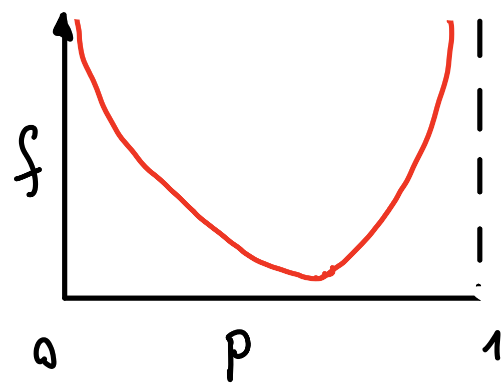


Some Problems

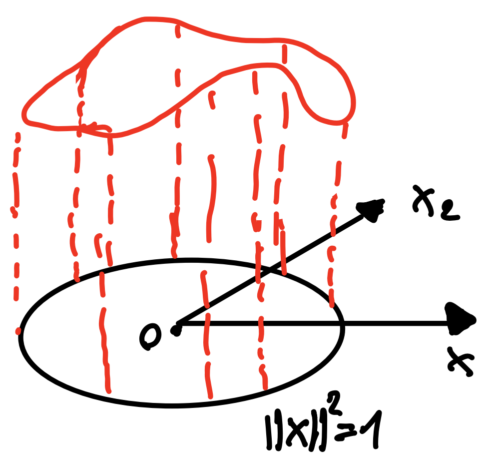
$$\min_{x>0} f(x)$$



$$\min_{0<p<1} f(p)$$

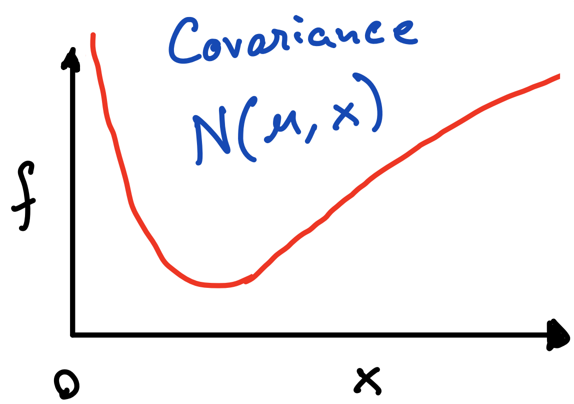


$$\min_{\|x\|^2=1} x^T A x$$

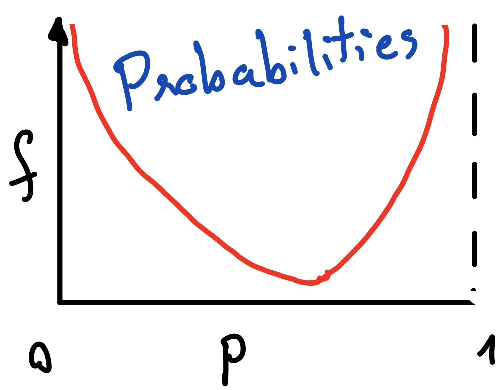


Some Problems

$$\min_{x>0} f(x)$$

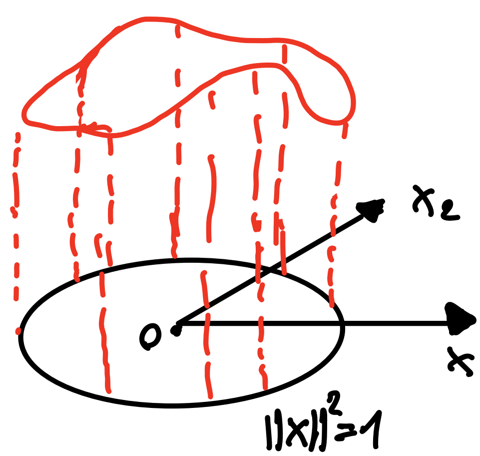


$$\min_{0<p<1} f(p)$$

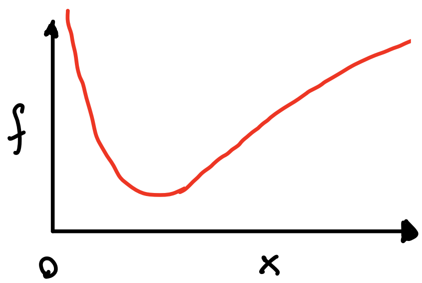


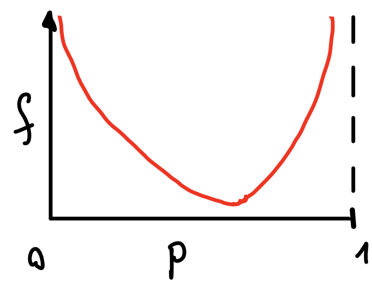
$$\min_{\|x\|^2=1} x^T A x$$

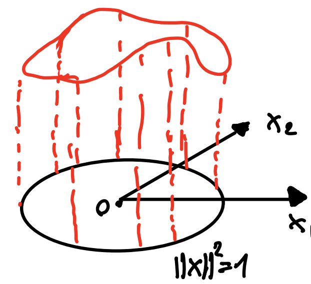
Eigenvectors



Some Problems



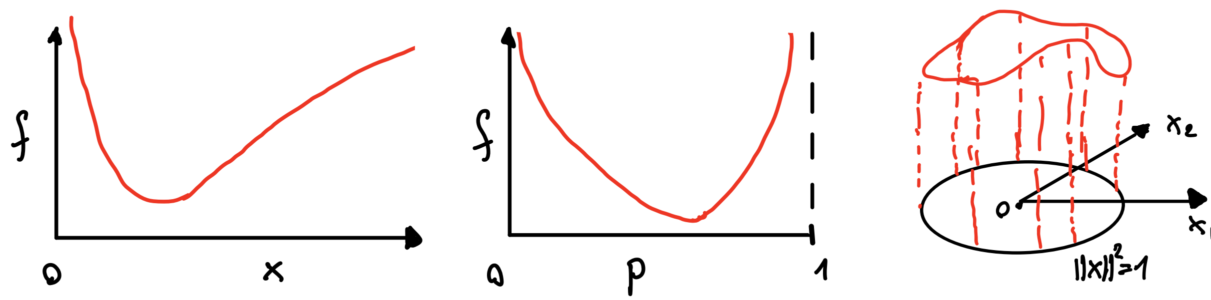




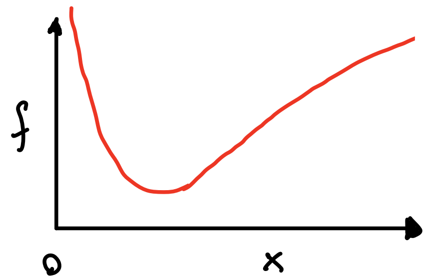
Not smooth Not convex Constrained

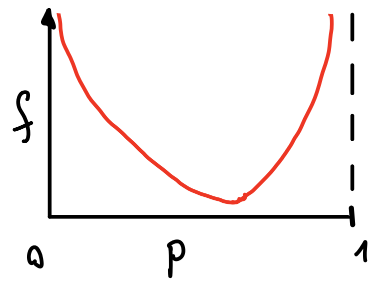
(but solvable)

Typical Solutions



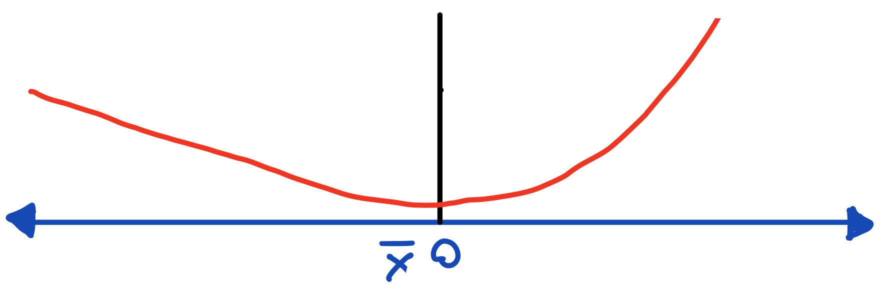
Typical Solutions



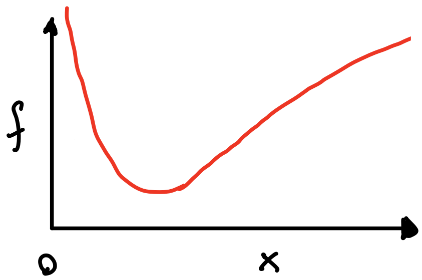


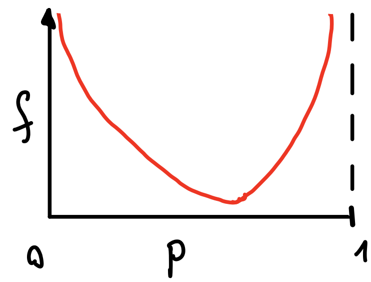
Change Parametrization

$$x \in (0, \infty) \;\text{-->}\; \bar{x} \in (-\infty, \infty) = \log(x)$$



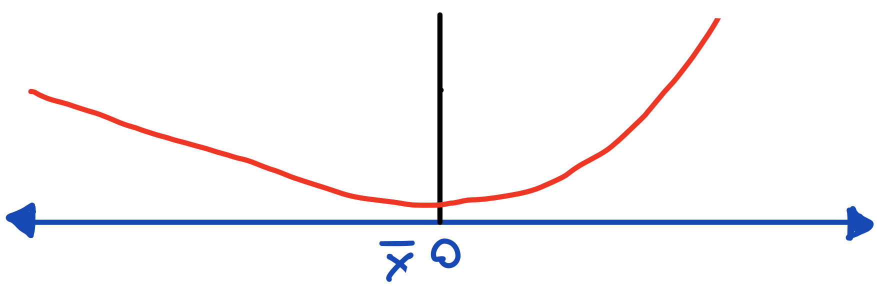
Typical Solutions



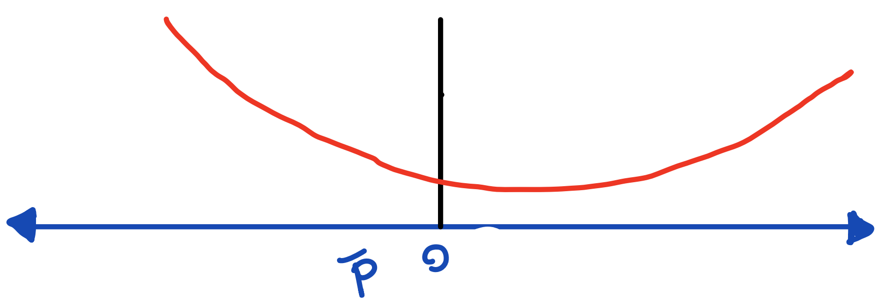


Change Parametrization

$$x \in (0, \infty) \to \bar{x} \in (-\infty, \infty) = \log(x)$$

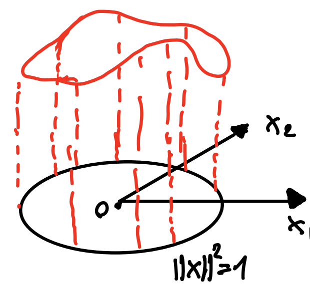


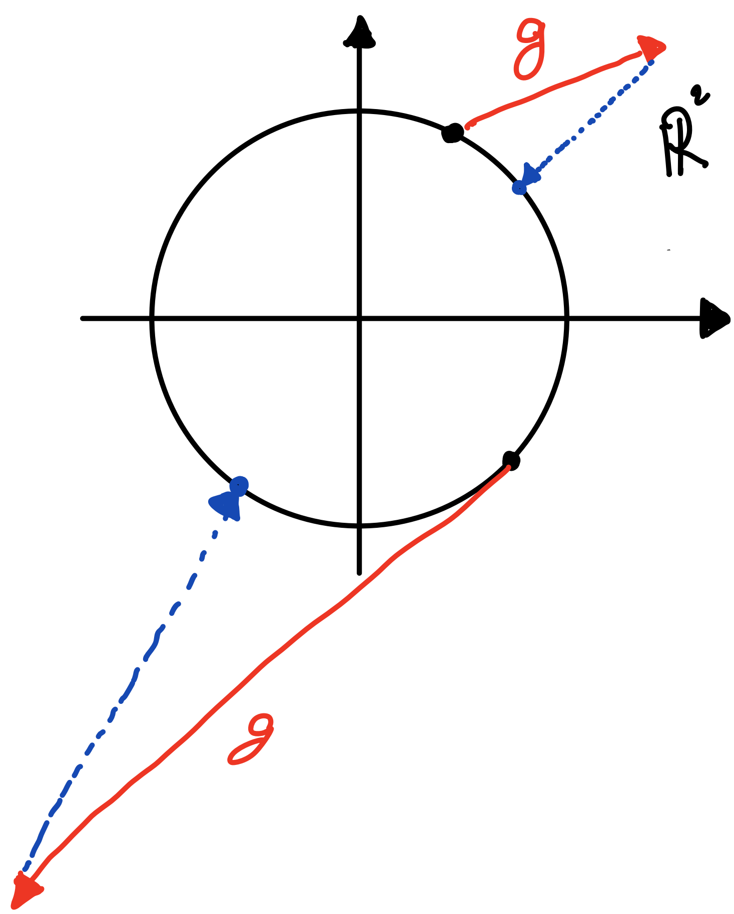
$$p \in (0, 1) \to \bar{p} \in (-\infty, \infty) = \sigma^{-1}(x)$$



Typical Solutions

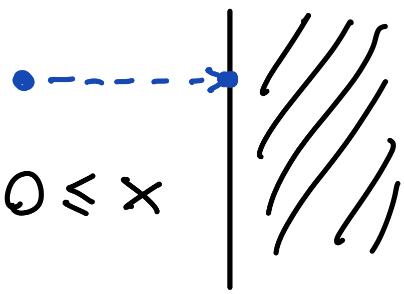
Projection



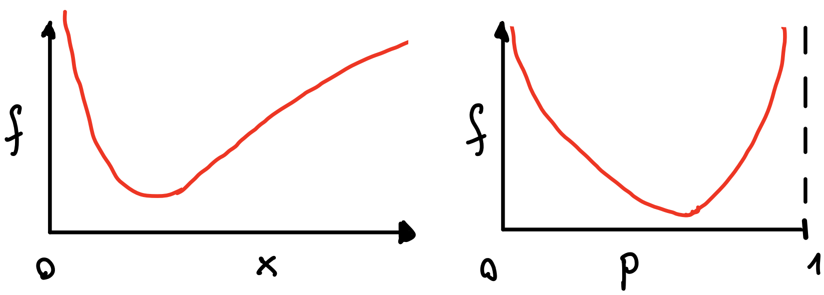


Some limitations

Projection



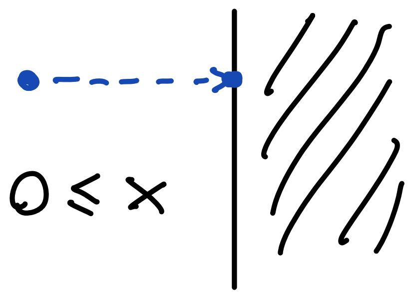
$0 \leq x$



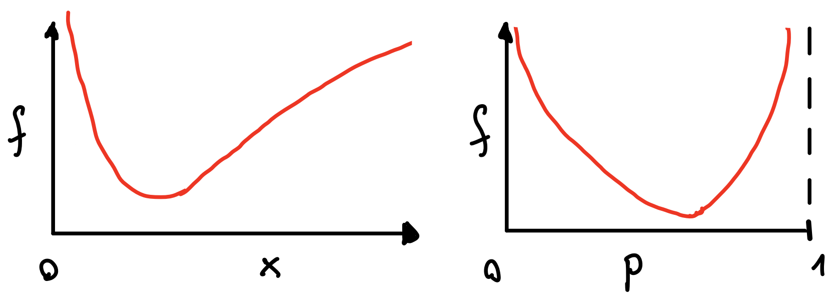
$0 < x$?

Some limitations

Projection

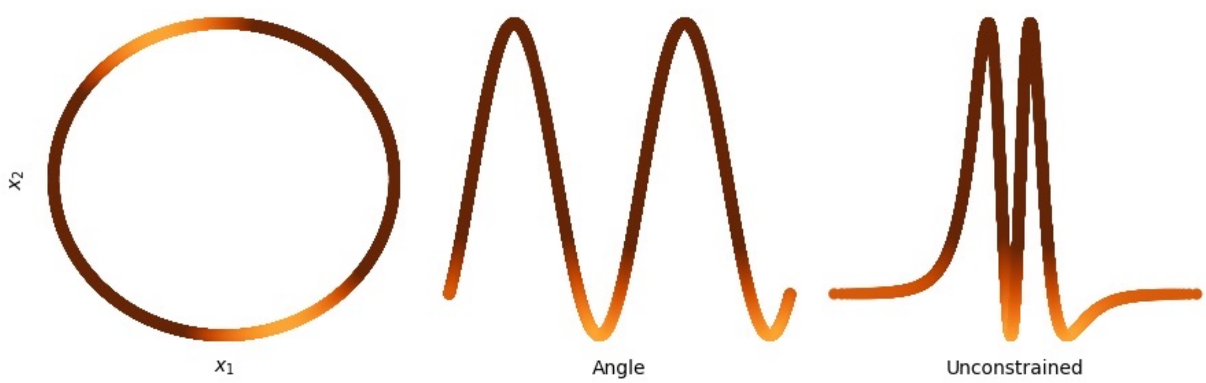


$0 \leq x$



$0 < x$?

Reparametrization



Tools <—> Problem

Projection

Reparam

Barrier functions

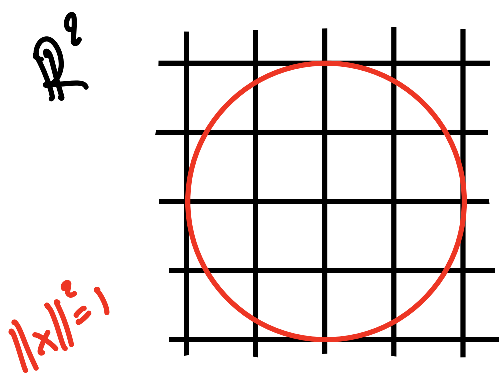
Lagrangian / dual

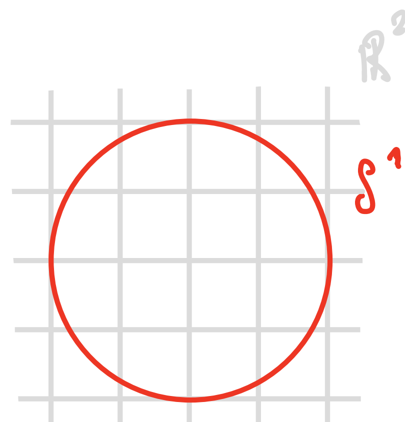
Mirror Descent

Tractable? Convex? Smooth?

Riemannian Opt "changes distances"

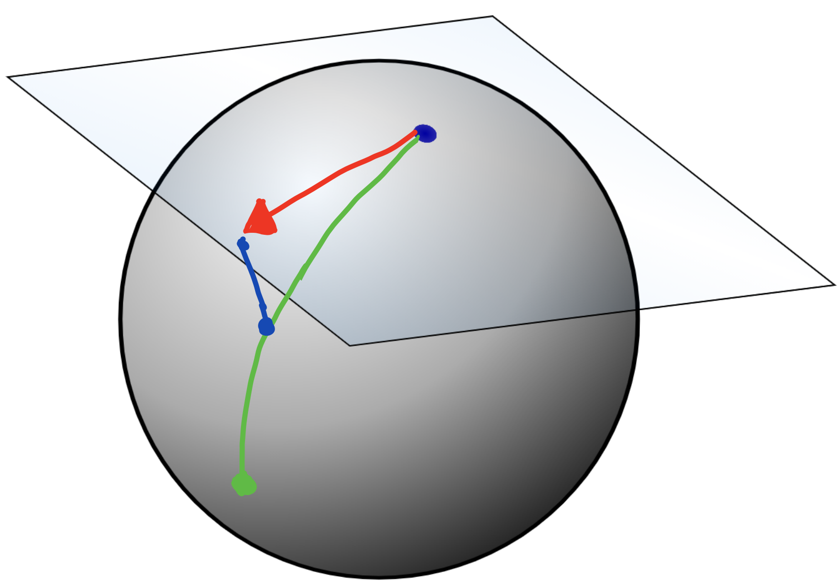
Euclidean vs Riemannian





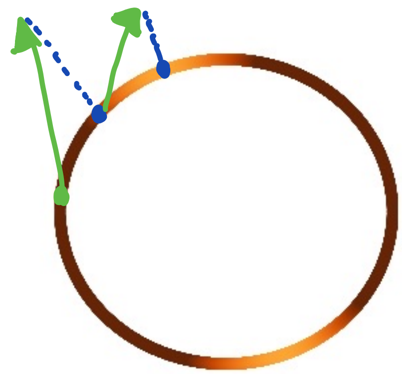
Constrained in $\mathbb{R}^2$

Unconstrained in $\mathcal{S}^1$

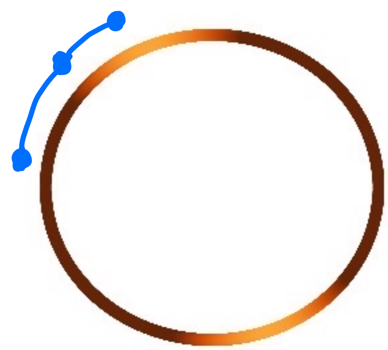


Riemannian ≠ Flows

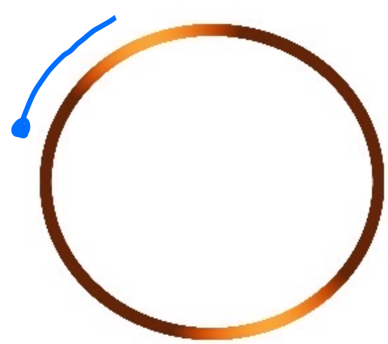
Euclidean
(+ proj)



Riemannian
GD

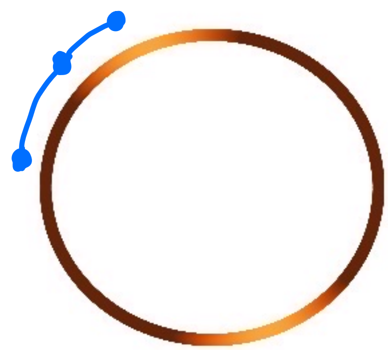


Flow



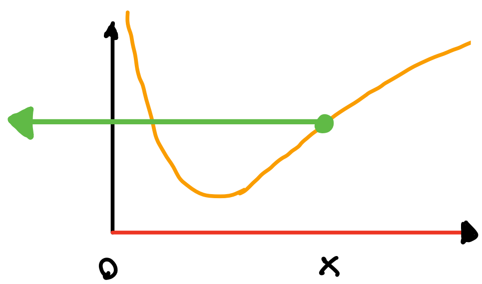
Riemannian GD

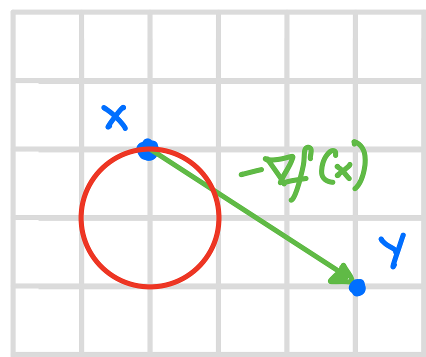
How?



Gradient descent is a lie*

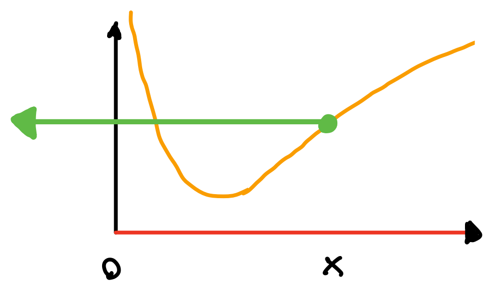
$$y = x - \nabla f(x)$$

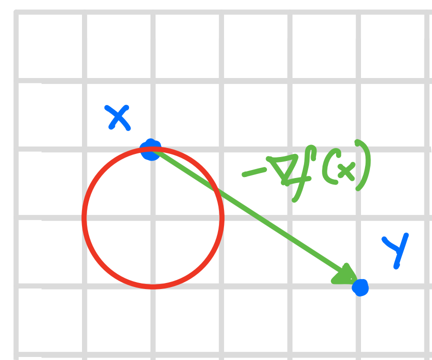




Gradient descent is a lie*

$$y = x - \nabla f(x)$$





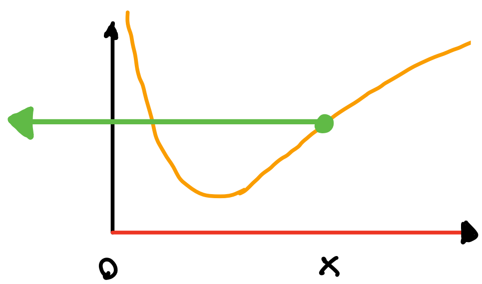
$x : [m]$ $\quad f : [m] \to [m^2]$

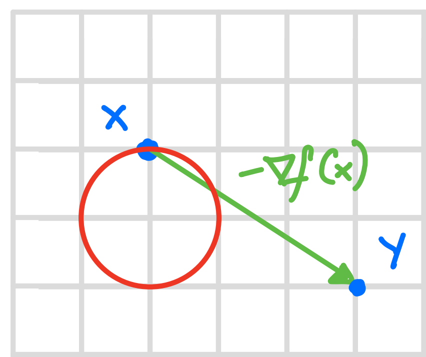
$f' = \frac{\Delta f}{\Delta x}$ $\quad \frac{[\quad]}{[\quad]} = [\quad]$

$y = [m] - [\quad]$

Gradient descent is a lie*

$$y = x - \nabla f(x)$$





$$x : \left[\frac{m}{s}\right] \qquad f : \left[\frac{m}{s}\right] \rightarrow [m]$$

$$f' = \frac{\Delta f}{\Delta x} \qquad \frac{[\quad]}{[\quad]} = [\quad]$$

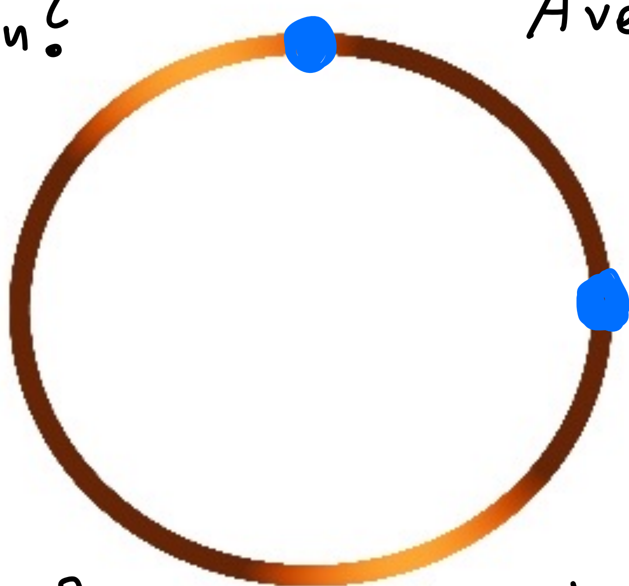
$$y = [m/s] - [\quad]$$

Riemannian Geometry

$$x \in S^1 \subset \mathbb{R}^2$$

Addition?

Average?

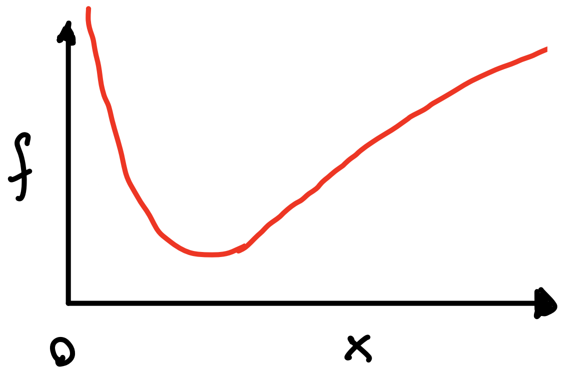


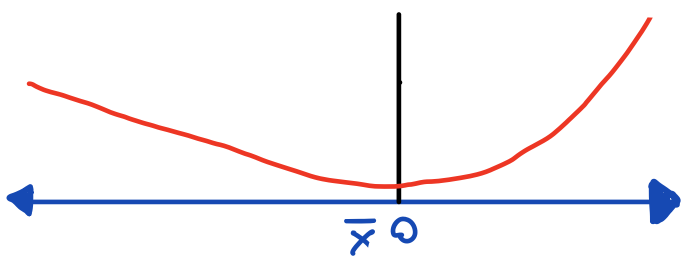
Length?

Norm?

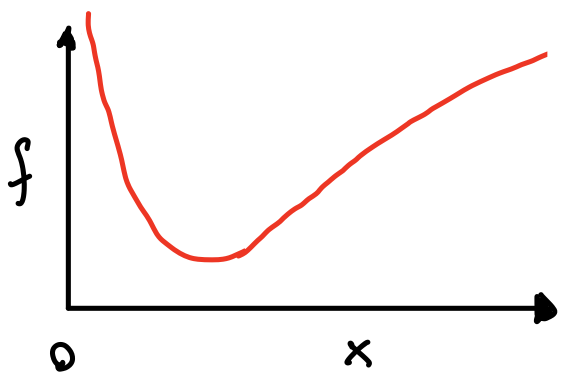
Inner product?

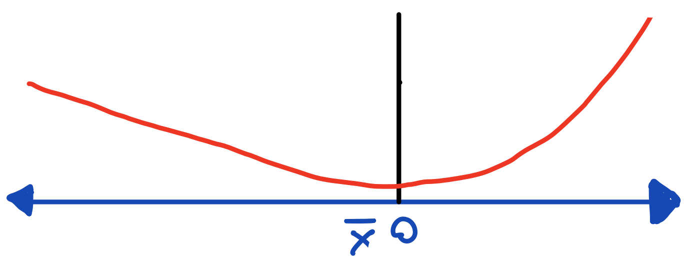
Distances

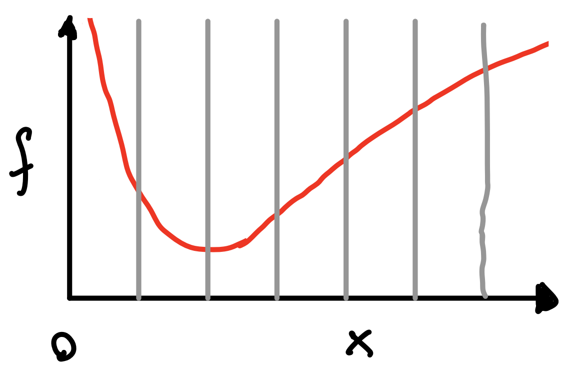


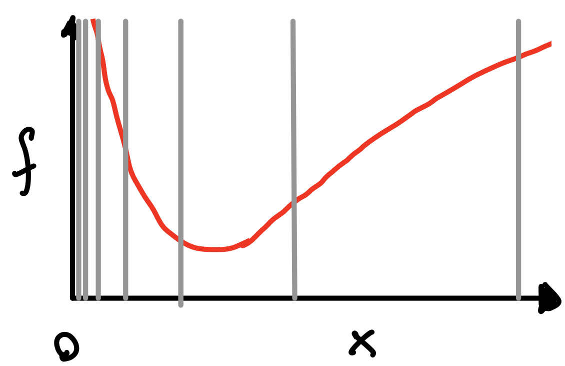


Distances





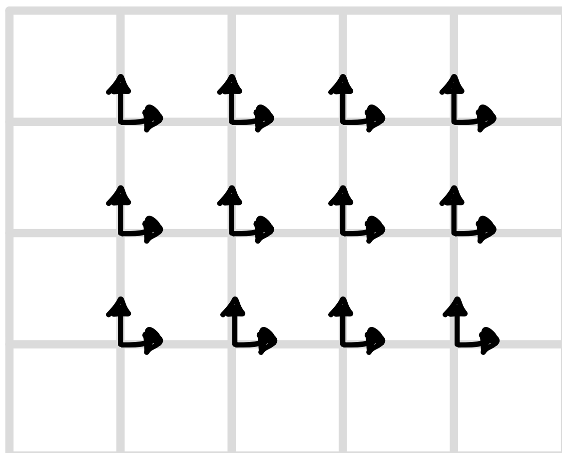




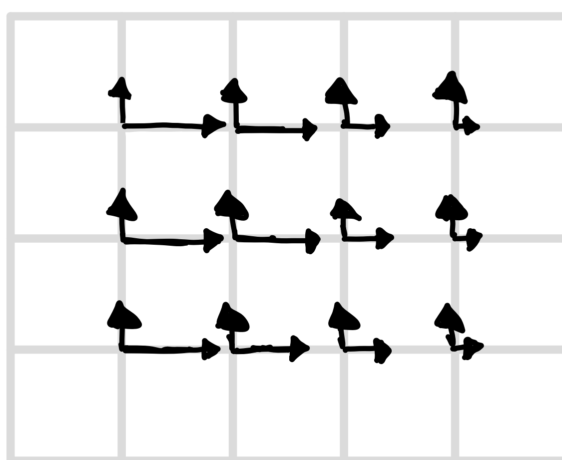
Metric

$$G(x) \in \mathbb{R}^{2\times 2}$$

Euclidean



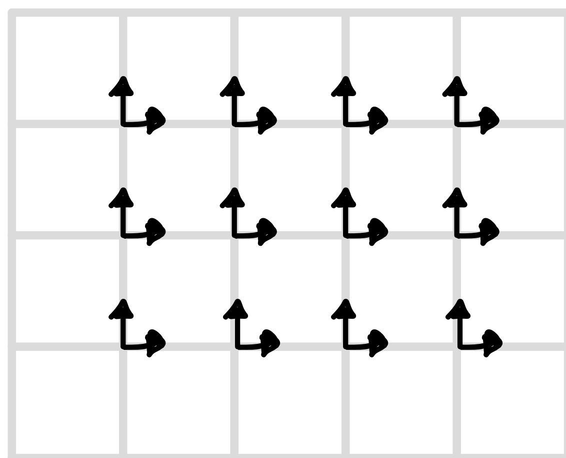
Something else



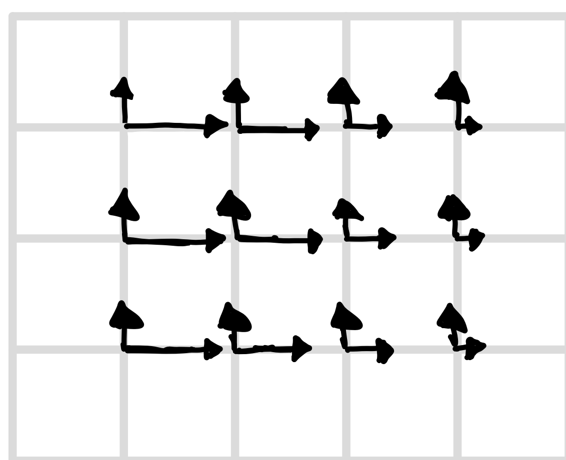
Metric

$$G(x) \in \mathbb{R}^{2\times 2}$$

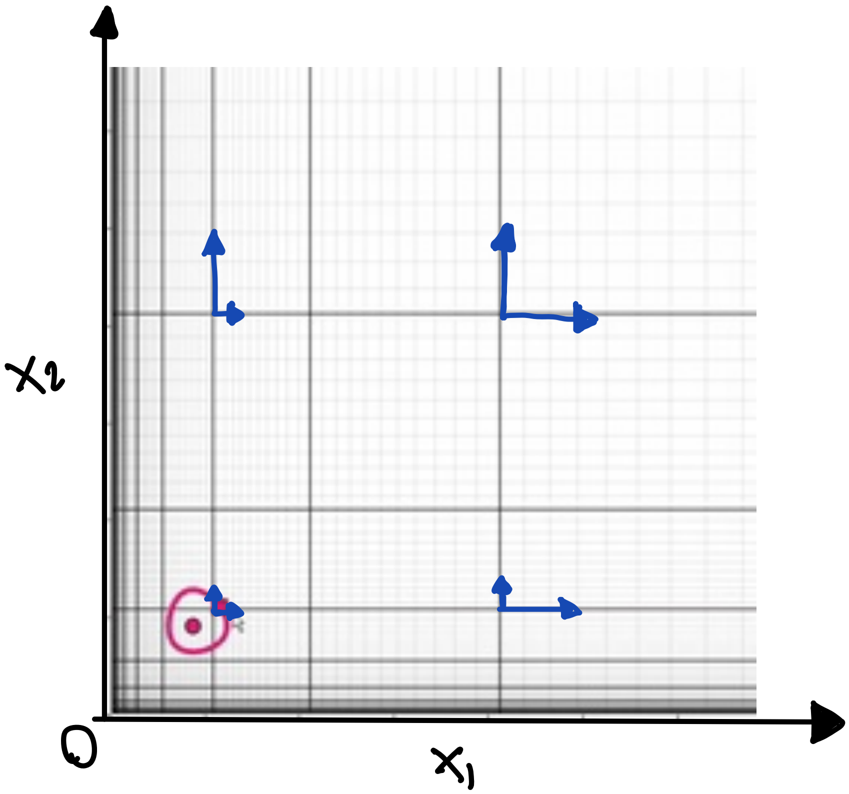
Euclidean



Something else

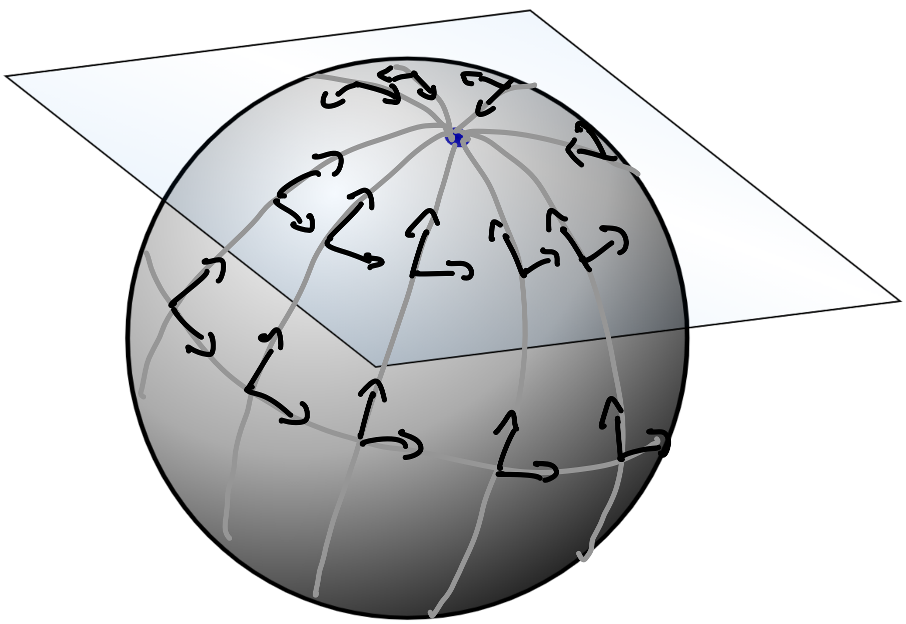


Metric

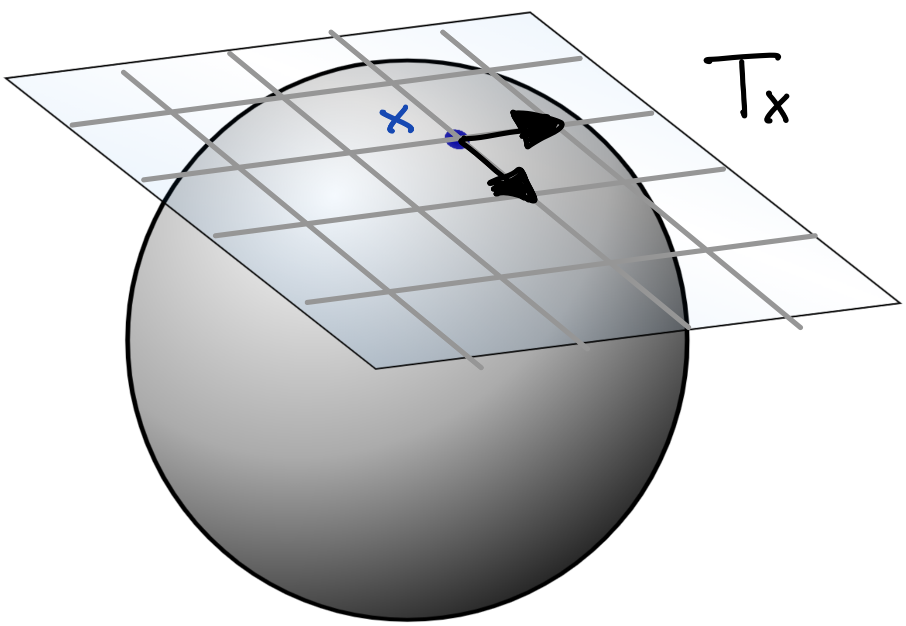


$$G(x) = \begin{bmatrix} \frac{1}{x_1^2} & 0 \\ 0 & \frac{1}{x_2^2} \end{bmatrix}$$

Tangent Space



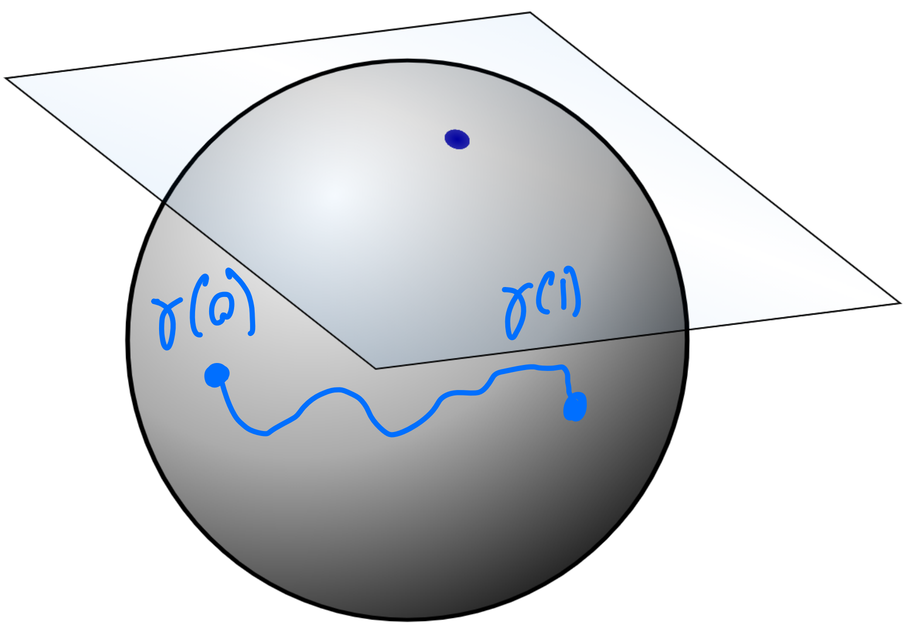
Tangent Space



$$\langle a, b \rangle_x = a^T G(x)\, b$$

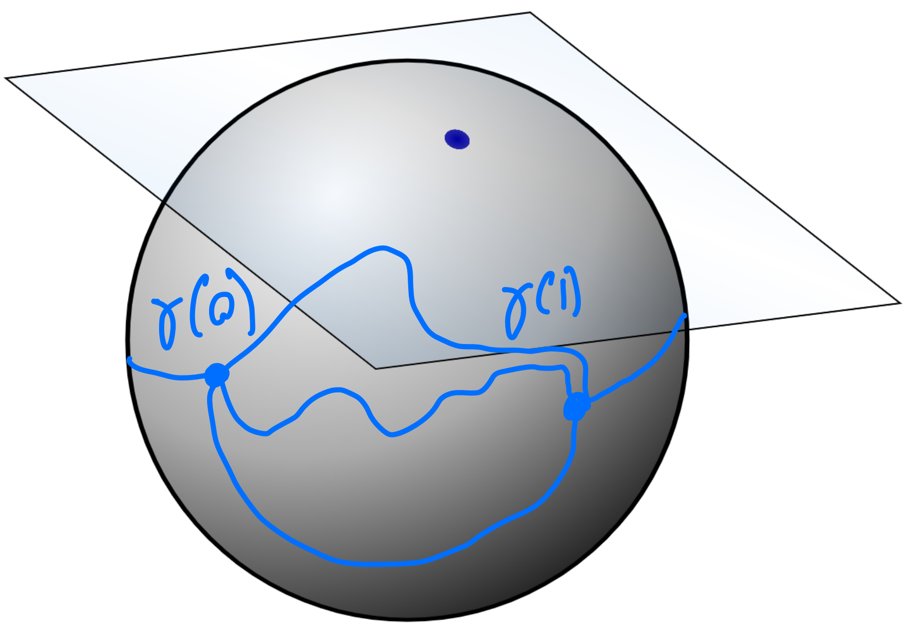
$$\|a\|_x = \sqrt{\langle a, a \rangle_x}$$

Paths

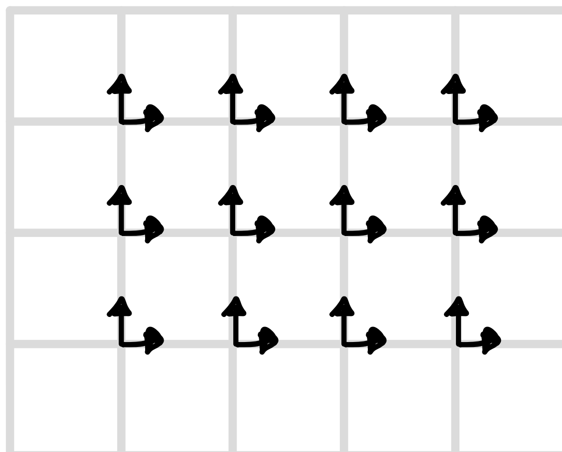


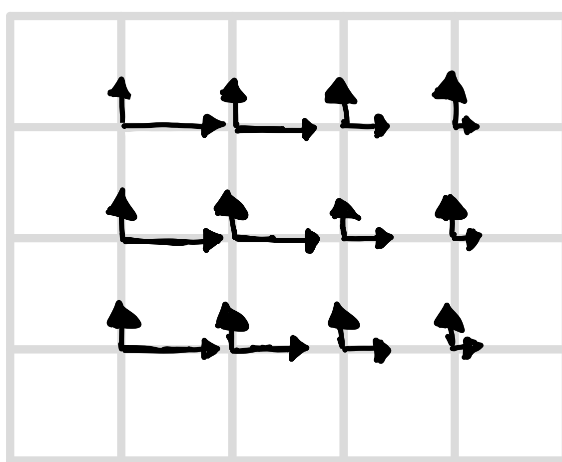
$\text{length}(\gamma) =$

Geodesic

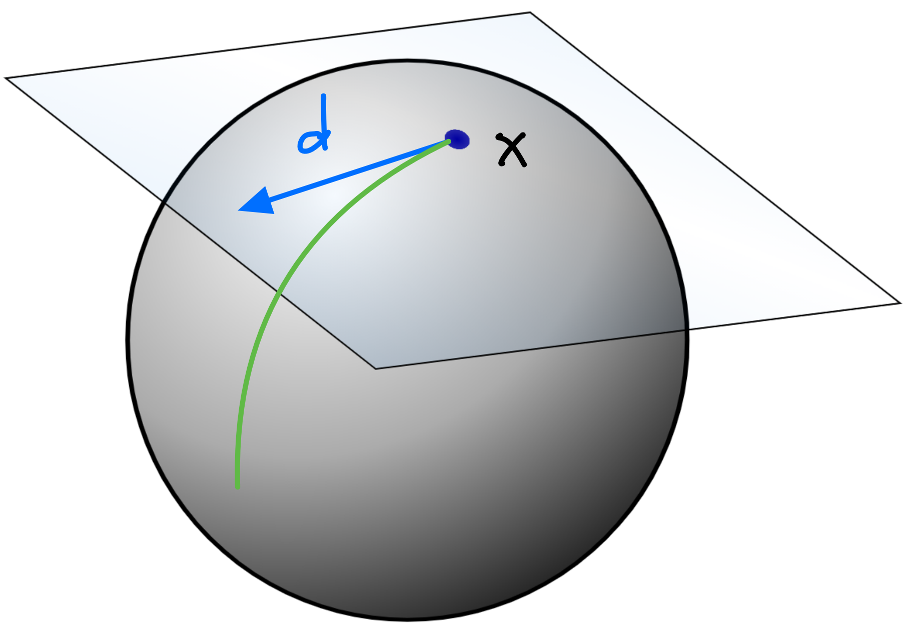


Geodesic





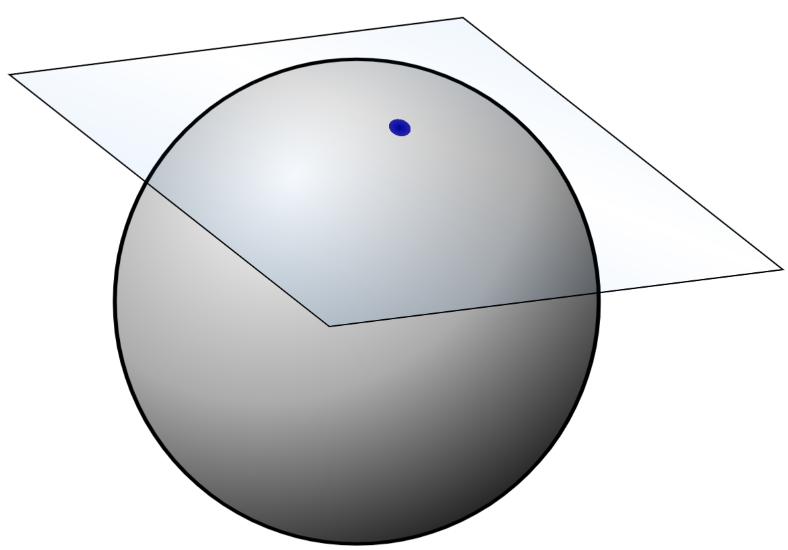
Geodesics



$\gamma(t)$

$$\gamma(0) = x$$
$$\dot{\gamma}(0) = d$$
$$\nabla_{\dot{\gamma}} \dot{\gamma}(t) = 0$$

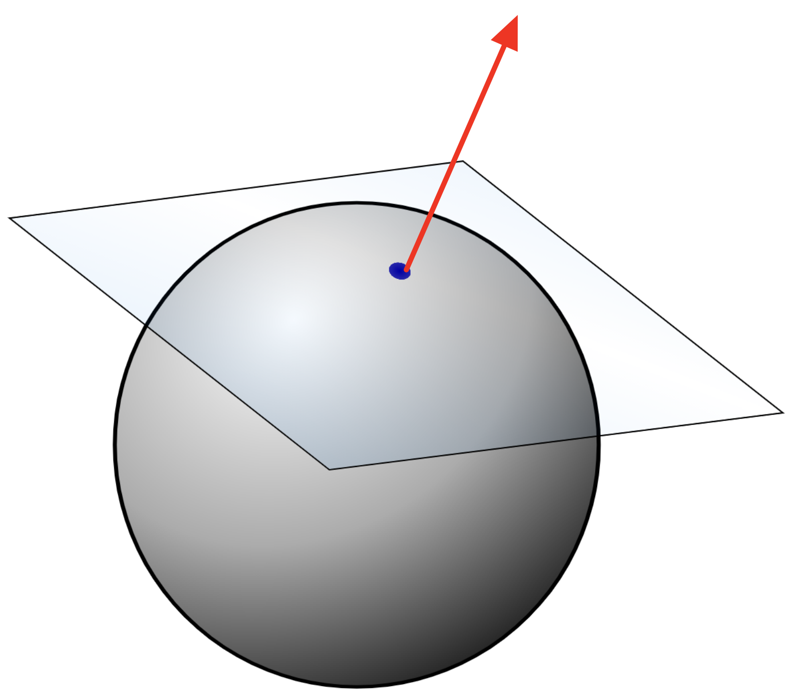
RGD on the sphere



RGD on the sphere

$$f(x) = \frac{1}{2} x^T A x$$

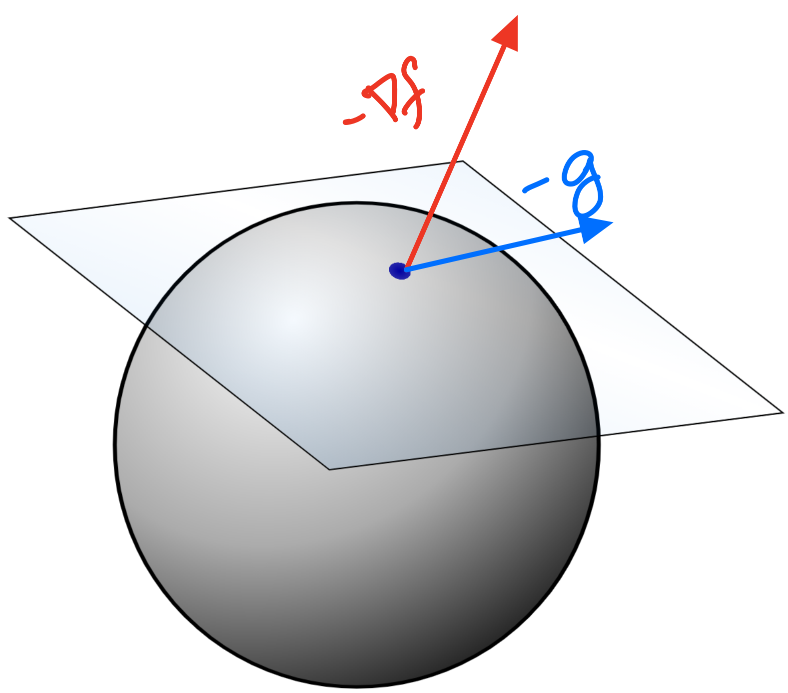
$$\nabla f(x) = A x$$



RGD on the sphere

$$f(x) = \frac{1}{2} x^T A x \qquad G(x) = I - x x^T$$

$$\nabla f(x) = A x \qquad g(x) = G(x)^{-1} \nabla f(x)$$

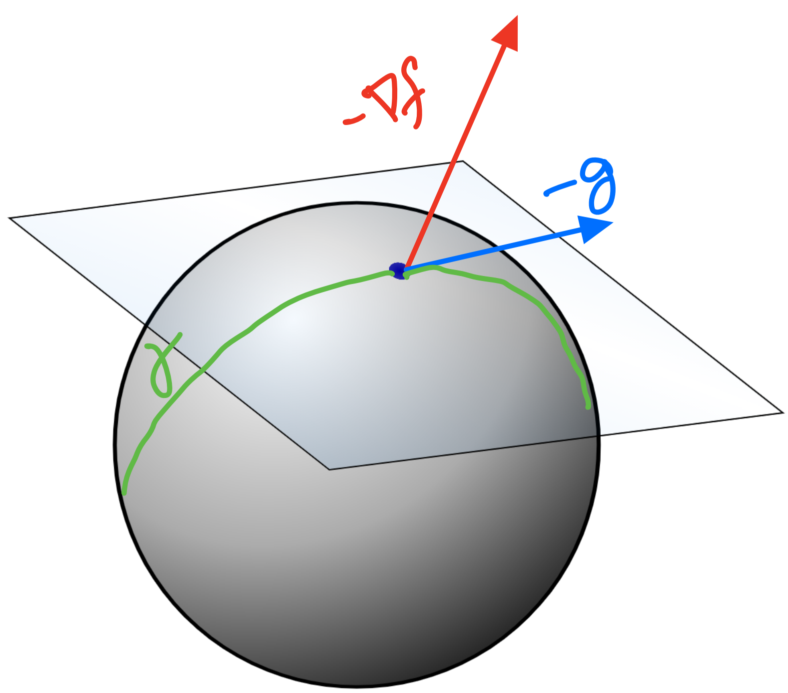


RGD on the sphere

$$f(x) = \frac{1}{2} x^T A x \qquad G(x) = I - x x^T$$

$$\nabla f(x) = A x \qquad g(x) = G(x)^{-1} \nabla f(x)$$

$$\gamma(t): \quad \gamma(0) = x, \quad \dot{\gamma}(0) = -g(x)$$

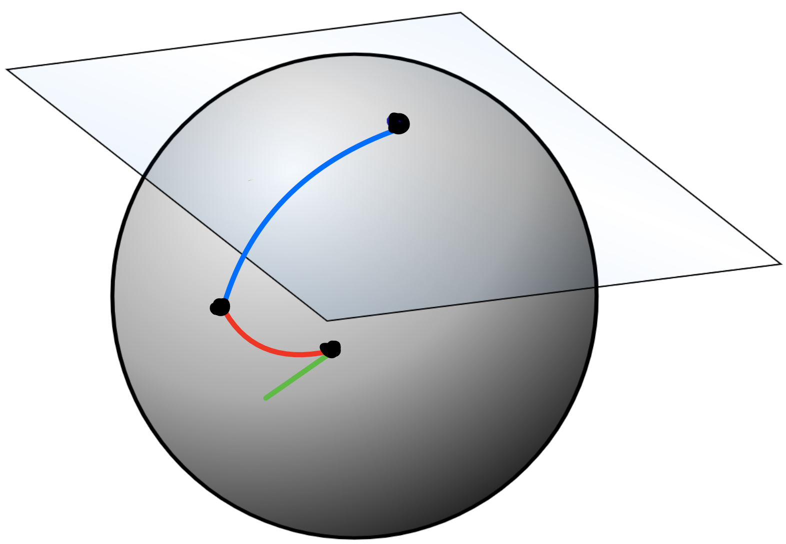


RGD on the sphere

$$f(x) = \frac{1}{2} x^T A x \qquad G(x) = I - x x^T$$

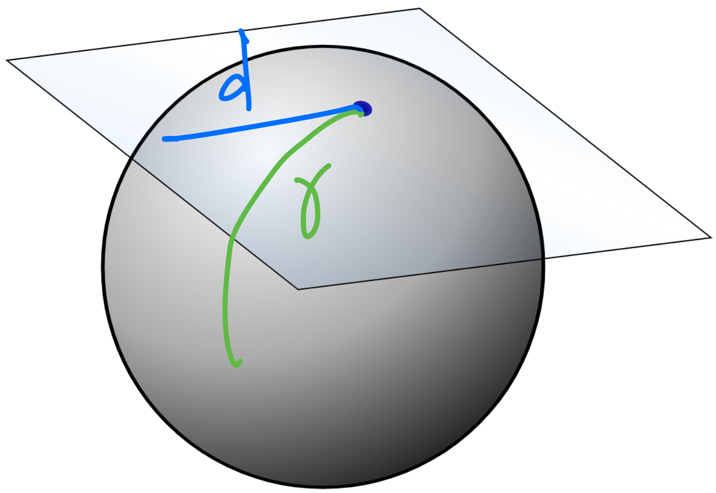
$$\nabla f(x) = A x \qquad g(x) = G(x)^{-1} \nabla f(x)$$

$$\gamma(t) = \cos(\|g\|_x t)\, x + \frac{\sin(\|g\|_x t)}{\|g\|_x} g$$



Quick Round

Solving DE



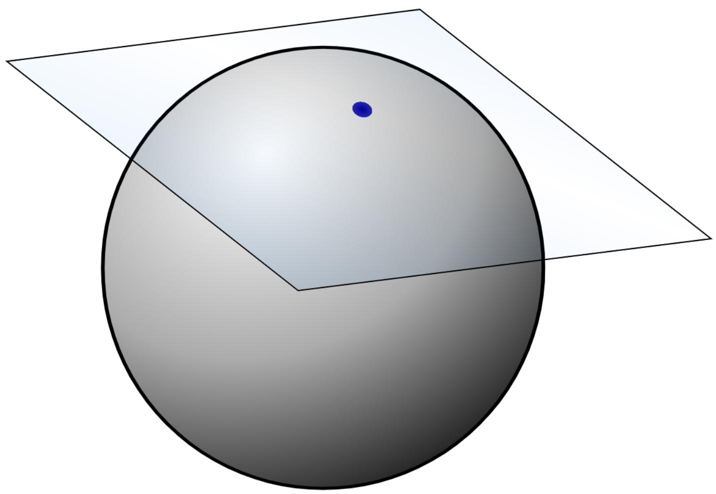
$$\gamma(0) = x$$

$$\dot\gamma(0) = d$$

$$\nabla_{\dot\gamma}\dot\gamma(t) = 0$$

Quick Round

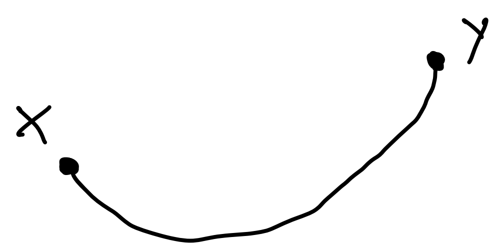
Retraction



Quick Round

Convexity

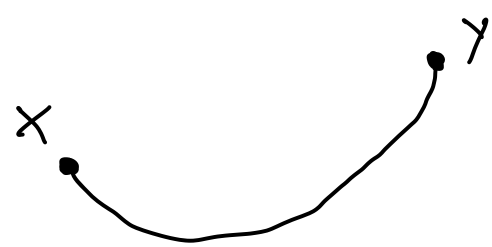
$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$



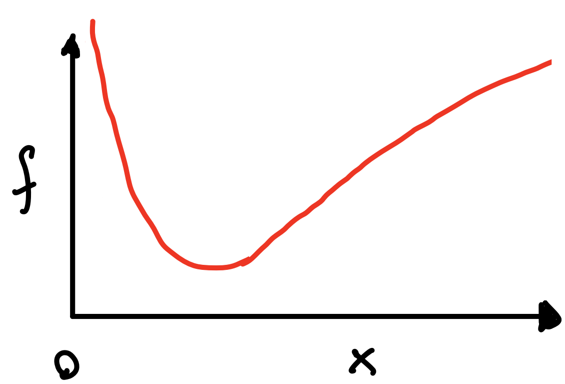
Quick Round

Convexity

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$



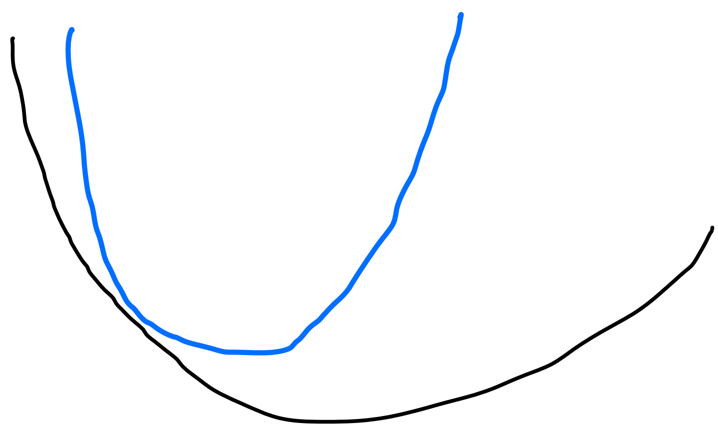
$$f\left(\gamma\left(\tfrac{1}{2}\right)\right) \leq \frac{1}{2}f(\gamma(0)) + \frac{1}{2}f(\gamma(1))$$



$$f\left(x^{1/2}y^{1/2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

Quick Round

Smoothness

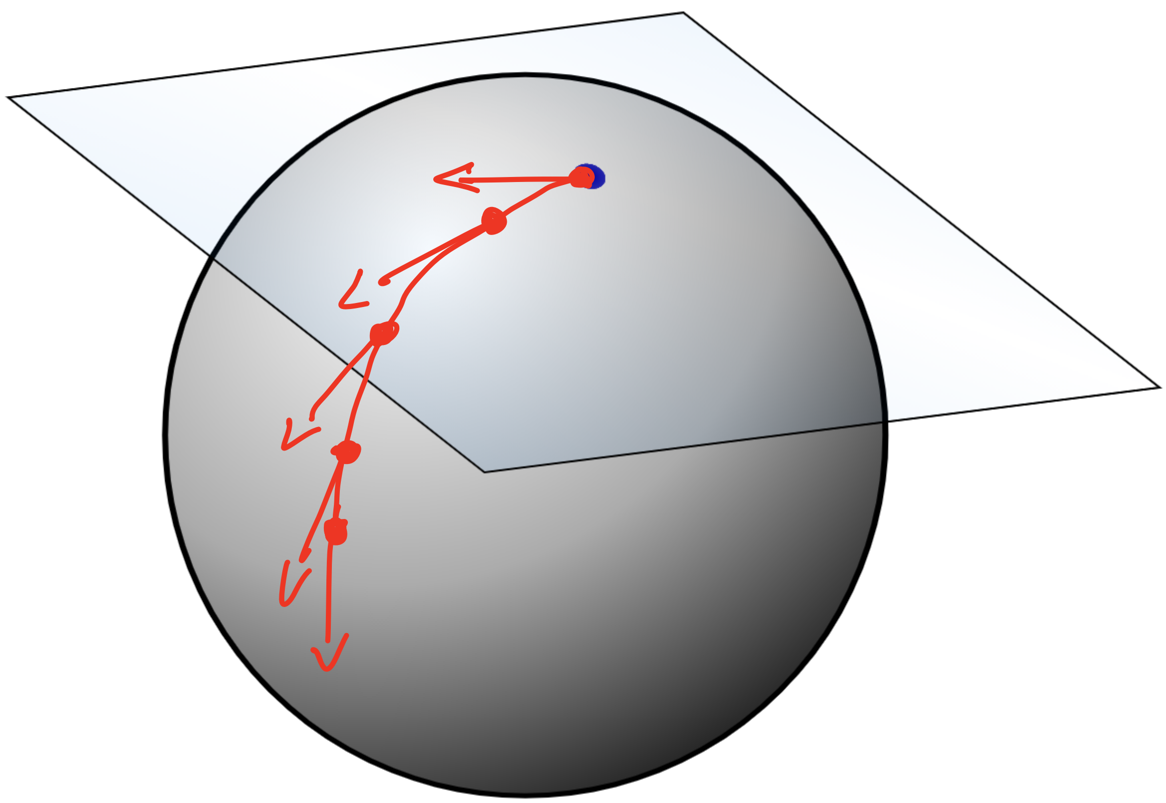


$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|$$

$$\|\nabla^2 f(x)\| \leq L$$

Quick Round

Parrallel Transport



Quick Round

Many Manifolds

$\{x > 0\}$ $\{\|x\|^2 = 1\}$

Quick Round

Many Manifolds

$\{x > 0\}$ $\{\|x\|^2 = 1\}$

$\{\Sigma \in \mathbb{R}^{d \times d} : \Sigma \succ 0\}$

$\{A \in \mathbb{R}^{d \times d} : \text{rank}(A) = k\}$

$\{Q \in \mathbb{R}^{d \times d} : QQ^T = I\}$

References

Books and monographs

An introduction to optimization on smooth manifolds
Nicolas Boumal

Riemannian Optimization and Its Applications
Hiroyuki Sato

Optimization Algorithms on Matrix Manifolds
Pierre-Antoine Absil, Robert Mahony, Rodolphe Sepulchre

Geodesic Convex Optimization: Differentiation on Manifolds, Geodesics, and Convexity
Nisheeth K. Vishnoi

Presentations

An Introduction to Geodesic Convexity
Nisheeth Vishnoi - watch?v=hJdcd1SR_tA

Covariance estimation and geodesic convexity
Ami Wiesel - watch?v=7-Yq3dpoUWM

Some non-convex optimization problems through a geometric lens
Suvrit Sra - watch?v=ys2XPPijoDA

Accelerated Gradient Methods on Riemannian Manifolds
Suvrit Sra - watch?v=-1BK7OHM6ME

Blogs

Optimization and Gradient Descent on Riemannian Manifolds
Agustinus Kristiadi - wiseodd.github.io/techblog