No. ○○○

Research Topic
Riemannian Optimization and Its Applications

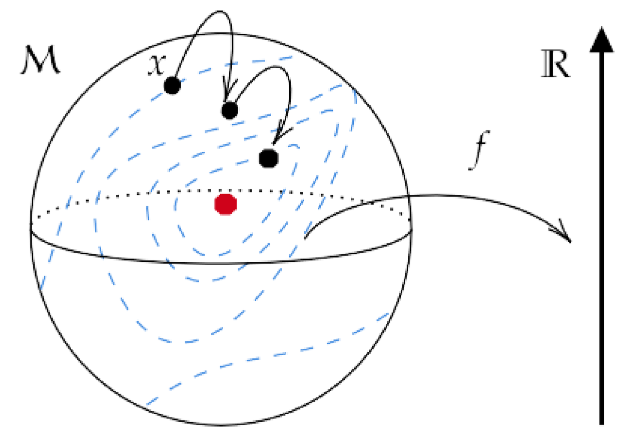
Affiliation: Graduate School of Systems and Information Engineering, University of Tsukuba.
Year: Doctor, 1st year　Name: 賴 志堅（ライ シケン）
Presentation Inquiries: s1920311@s.tsukuba.ac.jp

A Brief Introduction to Riemannian Optimization

Problem: Given $f(x): \mathcal{M} \to \mathbb{R}$, solve

$$\min_{x\in\mathcal{M}} f(x)$$

where $\mathcal{M}$ is a Riemannian manifold(リーマン多様体).



◆**Euclidean constrained problem ⇒ Riemannian unconstrained problem**

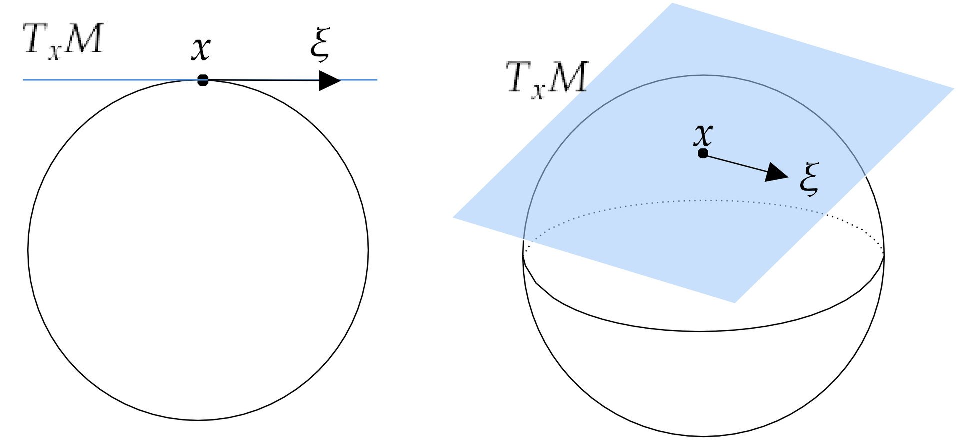
E.g. $\min_{x\in\mathbb{R}^n} f(x)$ s.t. $x^\top x = 1 \Rightarrow \min_{x\in\mathbb{S}^{n-1}:=\{x\in\mathbb{R}^n: x^\top x=1\}} f(x)$

Unit sphere, $\mathbb{S}^{n-1}$, is a Riemannian manifold.

- n=2, a circle;
- n=3, a sphere.

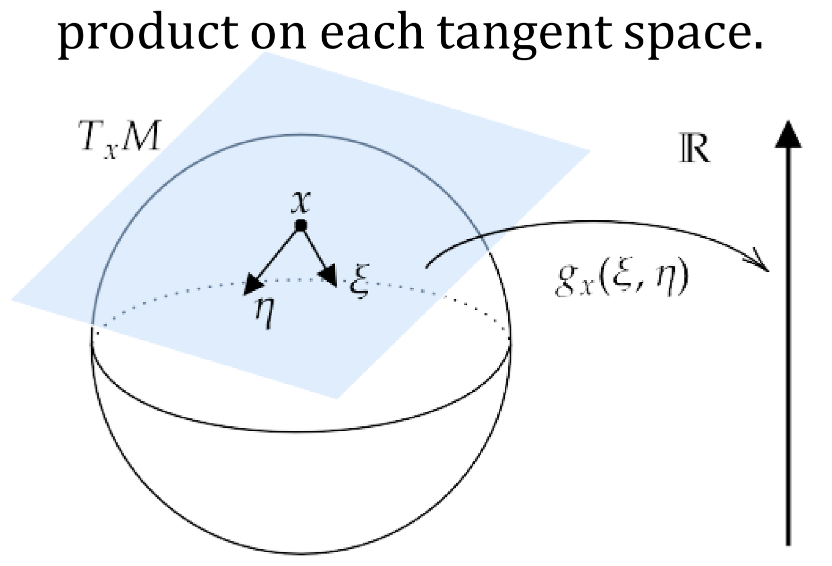
◆**Riemannian manifold = manifold(多様体) + Riemannian metric(リーマン計量).**

① Manifold is a set that can be locally linearized.



$T_x\mathcal{M}$ is tangent space at x.
$\xi \in T_x\mathcal{M}$ is tangent vector at x.

② A Riemannian metric, g, is an inner product on each tangent space.



Iterations on the Manifold

◆ **Euclidean Iterations:**

$$x_{k+1} = x_k + \alpha_k d_k$$

This iteration is implemented in numerous ways, e.g.:

- Steepest descent: $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$.
- Newton method: $x_{k+1} = x_k - \alpha_k[\nabla^2 f(x_k)]^{-1}\nabla f(x_k)$.

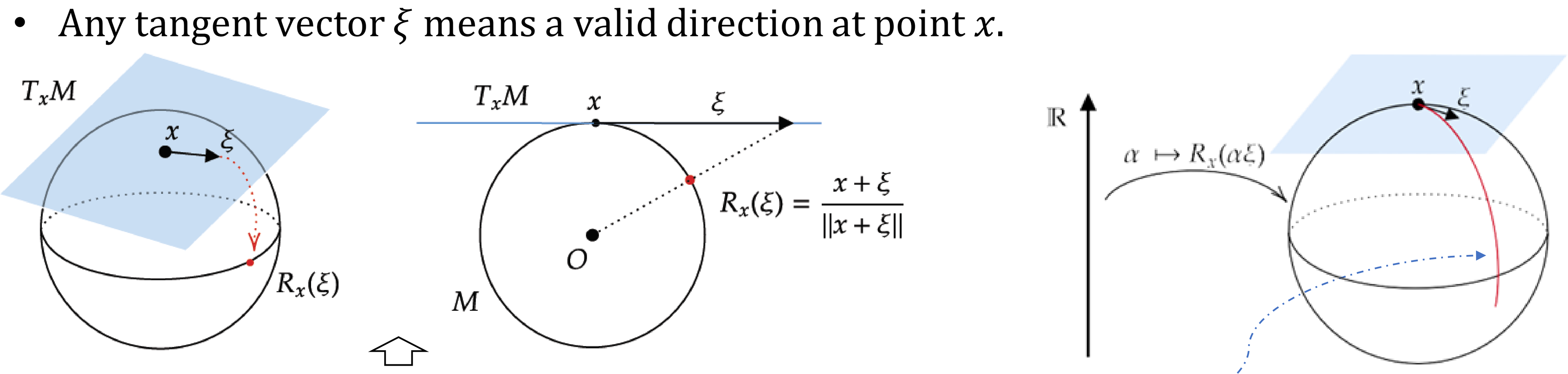
⇨

◆ **Riemannian Manifolds Provide:**

① directions and movement on manifold.
② Riemannian gradient and Hessian.

① **Moving on a manifold — retractions**

- Any tangent vector ξ means a valid direction at point x.

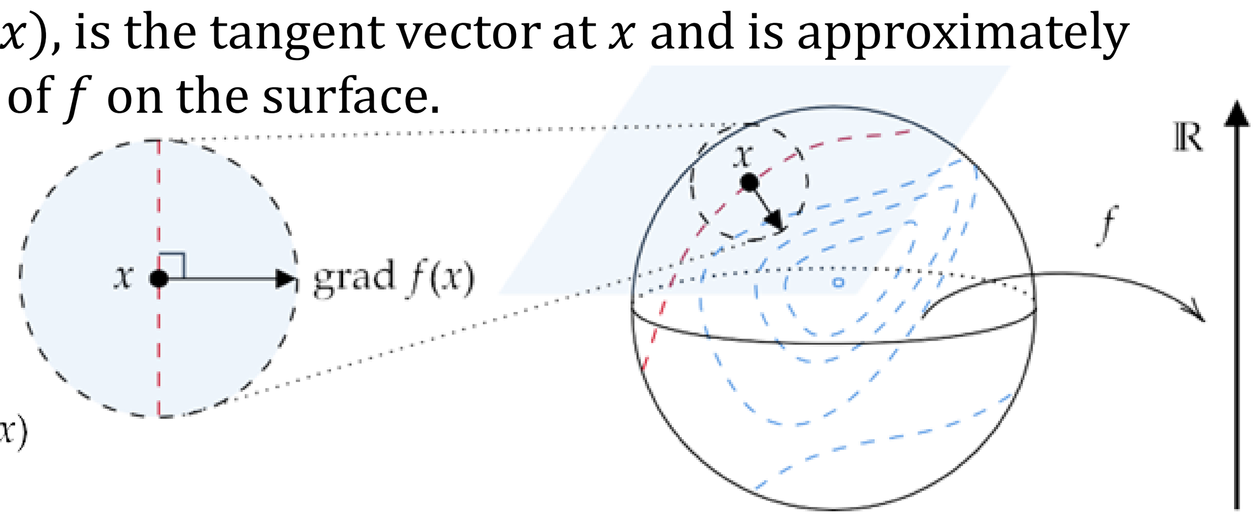


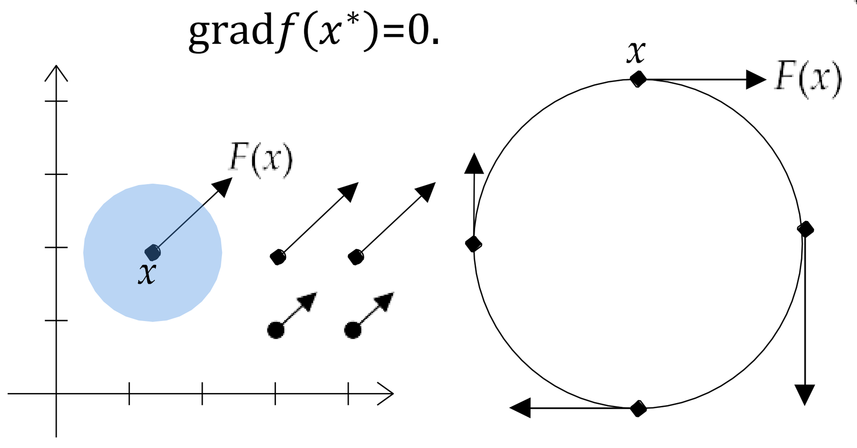
- A retraction R yields a map $R_x: T_x\mathcal{M} \to \mathcal{M}$ for any x.
- Given a tangent vector ξ at x, $\alpha \mapsto R_x(\alpha\xi)$ defines a curve in this direction.

② **Riemannian gradient — an intuitive definition**

- The Riemannian gradient, grad$f(x)$, is the tangent vector at x and is approximately perpendicular to the contour line of f on the surface.

a) $-$grad$f(x)$ is the direction of steepest descent at x;
b) If x^* is a local minimizer, then grad$f(x^*)$=0.





A vector field F on $\mathcal{M}$ is an assignment of a tangent vector to each point in M.

- grad$f$ is a special vector field on $\mathcal{M}$.

Applications

◆**Compute the single extreme eigenvalue or singular value.**

For a symmetric matrix $A \in \mathrm{Sym}(n)$, we have

the smallest eigenvalue of $A = \min_{x\in\mathbb{S}^{n-1}} x^T Ax$.

Similarly, for a matrix $M \in \mathbb{R}^{m\times n}$, we have

the largest singular value of $M = \max_{x\in\mathbb{S}^{m-1}, y\in\mathbb{S}^{n-1}} x^T My$.

□ Unit sphere manifold, $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n: \|x\|_2 = 1\}$.
□ $\mathbb{S}^{m-1}\times\mathbb{S}^{n-1}$ is a product manifold.

◆ **Sparse PCA .** We want to to find principal eigenvectors with few nonzero elements:

$$\min_{X\in\mathrm{St}(p,n)} -\mathrm{trace}(X^\top A^\top AX) + \rho\|X\|_1.$$

where $\|X\|_1 := \sum_{ij}|X_{ij}|$ and $\rho > 0$ is a parameter.

□ Stiefel manifold, $\mathrm{St}(p,n) := \{X \in \mathbb{R}^{n\times p} \mid X^T X = I_p\}$.
□ Grassmann manifold ---- Set of all p−dimensional subspaces of $\mathbb{R}^n$. Applications: dimension reduction problems.

◆ **Low-Rank Matrix Completion.** We want to recover a low-rank matrix M by

$$\min_X \ \mathrm{rank}(X) \quad \text{s.t.} \quad X_{ij} = M_{ij}, (i,j) \in \Omega.$$

If $\mathrm{rank}(M) = r$ is known, an alternative model is

$$\min_{X\in\mathrm{Fr}(m,n,r)} \sum_{(i,j)\in\Omega} (X_{ij} - M_{ij})^2.$$

□ Fixed rank manifold, $\mathrm{Fr}(m,n,r) := \{X \in \mathbb{R}^{m\times n}: \mathrm{rank}(X) = r\}$.

Libraries of Riemannian Optimization

Other manifolds:

- Oblique manifold, $\{X \in \mathbb{R}^{m\times n}: \mathrm{diag}(X^T X) = I_n\}$.
- Generalized Stiefel manifold, $\{X \in \mathbb{R}^{n\times p}: X^\top BX = I_p\}$ for some $B \succ 0$.
- Manifold of symmetric positive semidefinite, fixed-rank with unit diagonal, $\{X \in \mathbb{R}^{n\times n}: X = X^\top \succcurlyeq 0, \mathrm{rank}(X) = k, \mathrm{diag}(X) = 1\}$.
- And many more.

List of Riemannian methods (2002~):

- Steepest decent
- Newton
- trust region
- adaptive cubic overestimation
- conjugate gradient
- Quasi-Newton (BFGS)
- ADMM
- proximal gradient
- stochastic algorithms
- and many more.

Available solvers:

- Manopt (for Matlab, Python, Julia)
- McTorch (Riemannian optimization for deep learning)

Monographs:

- Optimization algorithms on matrix manifolds
- An introduction to optimization on smooth manifolds
- Riemannian Optimization and Its Applications

Survey:

- A brief introduction to manifold optimization

Current Research— Interior Point Methods for Manifolds

We try to solve **Riemannian constrained optimization problem**

$$\min_{x\in\mathbb{M}} f(x) \quad \text{s.t.} \quad h(x) = 0, \quad g(x) \geqslant 0,$$

where $f: \mathbb{M} \to \mathbb{R}$, $h: \mathbb{M} \to \mathbb{R}^l$, and $g: \mathbb{M} \to \mathbb{R}^m$.

Applications:

1. Nonnegative PCA: $\min_{X\in\mathbb{R}^{n\times p}} -\mathrm{tr}(X^\top AA^\top X)$ s.t. $X^\top X = I_p, X \geq 0$;
2. Subproblem of K-indicators model for Data Clustering;
3. Minimum Balanced Cut for Graph Bisection;
4. And many more.

Optimal conditions in Riemannian version: KKT conditions

Few existing Riemannian algorithms!

1. Augmented Lagrangian Method .
2. Exact Penalty Method .
3. Sequential Quadratic Method
4. Why not Interior Point Method?

This work was supported by JST SPRING, Grant Number JPMJSP2124.