Riemannian Interior Point Methods for Manifold Optimization

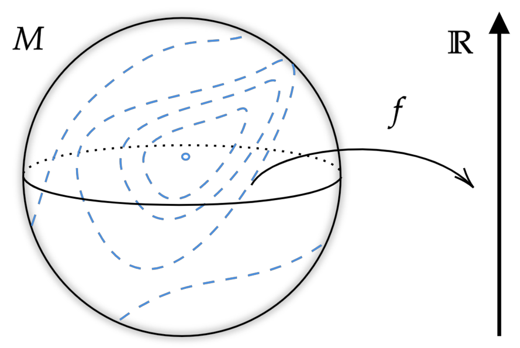
Year: Doctor, 2nd year
Name: ZHIJIAN LAI
Presentation Inquiries: s2130117@s.tsukuba.ac.jp
Affiliation: Graduate School of Systems and Information Engineering, University of Tsukuba.

A Brief Introduction to Riemannian Optimization

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

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

where $\mathcal{M}$ is a Riemannian manifold.



Constrained ⇒ Unconstrained
For example, consider

$$\text{minimize}_{x \in \mathbb{R}^n} f(x) \text{ subject to } x^\top x = 1,$$

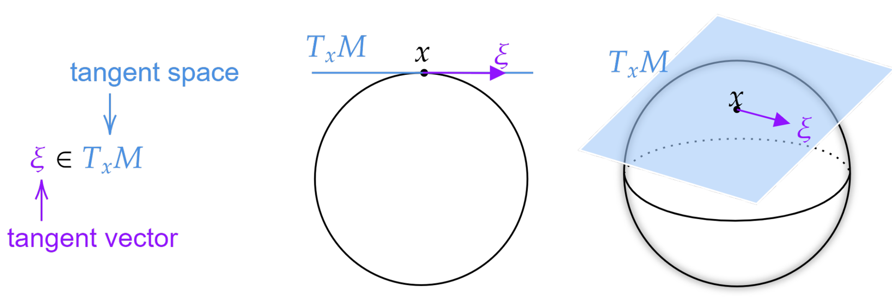
or,

$$\text{minimize}_{x \in \mathbb{S}^{n-1}} f(x),$$

where the unit sphere, $\mathbb{S}^{n-1} := \left\{x \in \mathbb{R}^n : x^\top x = 1\right\}$, is a Riemannian manifold.

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

Riemannian manifold is a space that can be locally linearized.



Iterations on the Euclidean Space
- x_k : current point.
- α_k : step size.
- d_k : direction vector.

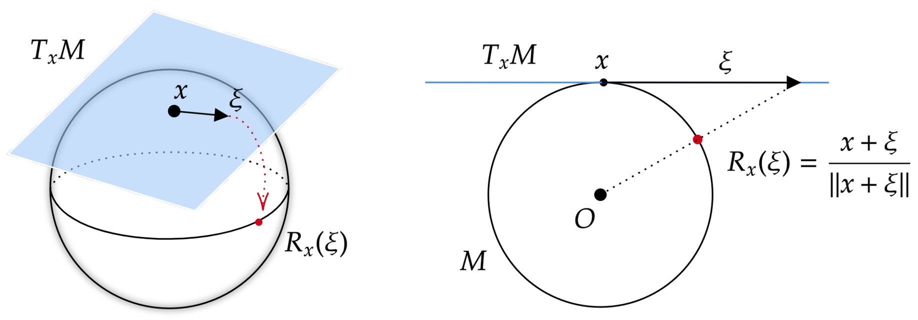
Next point is computed by

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

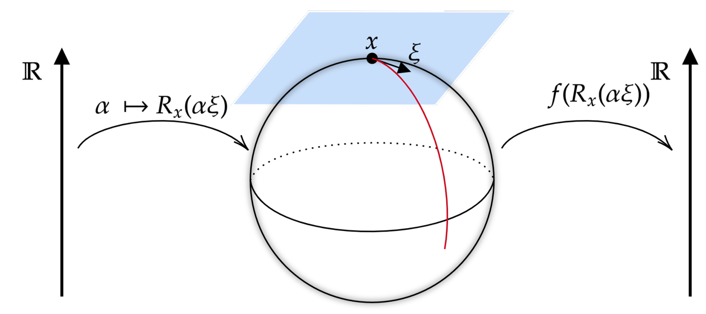
This iteration is implemented in numerous ways, e.g.,
- Steepest descent: $d_k = -\nabla f(x_k)$.
- Newton method: $d_k = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$.

Iterations on the Manifold

Moving on a manifold A retraction R yields a map $R_x : T_x\mathcal{M} \to \mathcal{M}$ for any x.

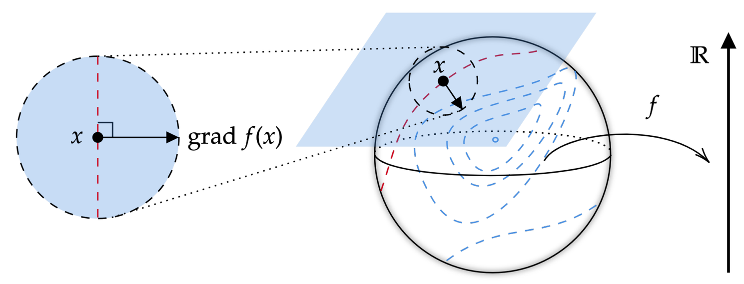


Given a tangent vector ξ at $x, \alpha \mapsto R_x(\alpha\xi)$ defines a curve in this direction.

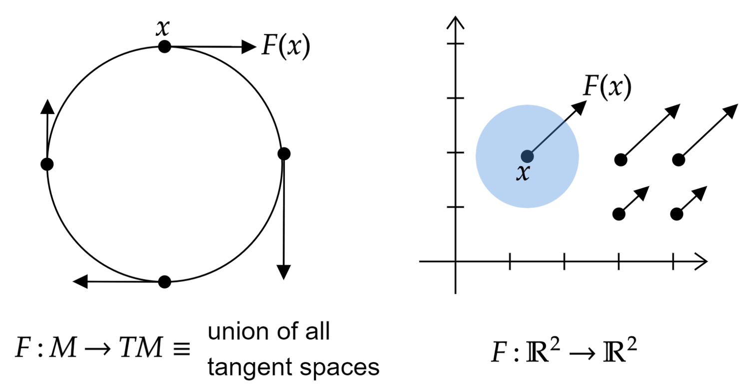


Riemannian Gradient , grad $f(x)$, is the tangent vector at x and is approximately perpendicular to the contour line of f on the surface.

- $-$ grad $f(x)$ is the direction of steepest descent at x;
- 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 $\mathcal{M}$. Thus, grad f is a special vector field on $\mathcal{M}$.



Application

Extreme eigenvalue or singular value
For a symmetric matrix $A \in \text{Sym}(n)$, we have the smallest eigenvalue of A which is equal to

$$\min_{x \in \mathbb{S}^{n-1}} x^T A x.$$

Similarly, for a matrix $M \in \mathbb{R}^{m \times n}$, we have the largest singular value of M which is equal to

$$\max_{x \in \mathbb{S}^{m-1}, y \in \mathbb{S}^{n-1}} x^T M y.$$

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:

$$\min_{x \in \mathrm{S}(p,n)} -\operatorname{trace}\left(X^\top A^\top A X\right) + \rho\|X\|_1.$$

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

$$\mathrm{St}(p, n) := \left\{X \in \mathbb{R}^{n \times p} \mid X^T X = I_p\right\}.$$

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

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

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

$$\min_{X \in \mathcal{M}_r} \sum_{(i,j) \in \Omega} (X_{ij} - M_{ij})^2.$$

Fixed rank manifold,

$$\mathcal{M}_r := \left\{X \in \mathbb{R}^{m \times n} : \operatorname{rank}(X) = r\right\}.$$

Current Research— Interior Point Methods for Manifolds

We extend the primal-dual interior point method from the Euclidean setting to the Riemannian one. Our method, named the Riemannian interior point method (RIPM), solves the Riemannian constrained optimization problems:

$$\begin{aligned} \min_{x \in \mathbb{M}} \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0, \text{ and } g(x) \le 0, \end{aligned} \tag{RCOP}$$

where $\mathbb{M}$ is a d-dimensional Riemannian manifold, $f : \mathbb{M} \to \mathbb{R}, h : \mathbb{M} \to \mathbb{R}^l$, and $g : \mathbb{M} \to \mathbb{R}^m$ are smooth functions. Such problem has wide applications, e.g.,

1. Nonnegative PCA: $\min_{X \in \mathbb{R}^{n \times p}} -\operatorname{tr}\left(X^\top A A^\top X\right)$ s.t. $X^\top X = I_p, X \ge 0$;
2. Subproblem of K-indicators model for Data Clustering;
3. Minimum Balanced Cut for Graph Bisection;

Riemannian optimality conditions:

- KKT conditions; Second-order conditions [Yang et al., 2014];
- More constraint qualifications (CQ) [Bergmann and Herzog, 2019];
- Sequential optimality conditions [Yamakawa and Sato, 2022].

Riemannian algorithms:

1. Augmented Lagrangian Method [Liu and Boumal, 2020, Yamakawa and Sato, 2022];
2. Exact Penalty Method [Liu and Boumal, 2020];
3. Sequential Quadratic Programming Method [Schiela and Ortiz, 2020, Obara et al., 2022].
4. In this talk, we consider Interior Point Method.

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