A Series of Talks on Riemannian Optimization

Introduction and Some Basics

Wen Huang

Xiamen University

September 15, 2025

Hunan University

Content

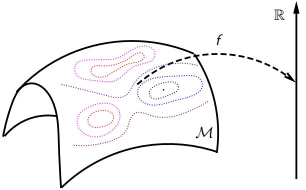
1. Introduction
   - Riemannian optimization;
   - Applications;
   - Smooth optimization framework;
   - Research foci of Riemannian optimization;
2. Embedded Submanifold
   - Definition;
   - Tangent space;
   - Smooth maps;
   - Vector fields and tangent bundle;
   - Retraction;
   - Riemannian Metric;
   - A Riemannian steepest descent algorithm;

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.

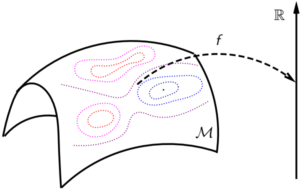


Riemannian Optimization

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

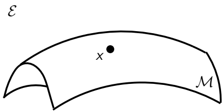
$$\min_{x \in \mathcal{M}} f(x)$$

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

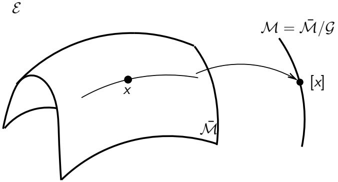


Two kinds of commonly-encountered manifolds

Embedded submanifold of a Euclidean space



Quotient manifold from an embedded submanifold

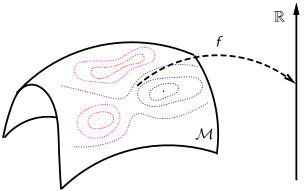


Riemannian Optimization

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

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

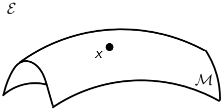
where $\mathcal{M}$ is a Riemannian manifold.



Examples:

- Sphere: $\{x \in \mathbb{R}^n \mid \|x\| = 1\}$;
- Stiefel manifold: $\mathrm{St}(p, n) = \{X \in \mathbb{R}^{n \times p} \mid X^T X = I_p\}$;
- Fixed rank: $\mathbb{R}_r^{m \times n} = \{X \in \mathbb{R}^{m \times n} : \mathrm{rank}(X) = r\}$;
- etc;

Embedded submanifold of a Euclidean space

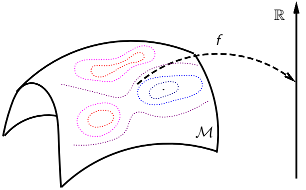


Riemannian Optimization

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

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

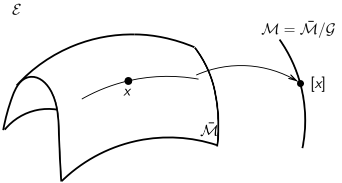
where $\mathcal{M}$ is a Riemannian manifold.



Examples:

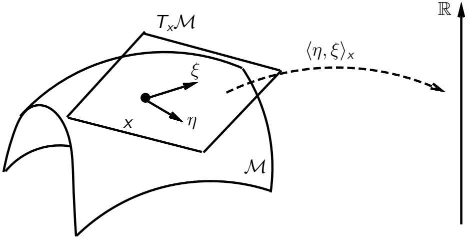
- Grassmann manifold:
  the set of p dimensional linear spaces in $\mathbb{R}^n$
  $\mathrm{Gr}(p, n) = \mathrm{St}(p, n)/\mathcal{O}_p$;
- Shape space;
- etc;

Quotient manifold from an embedded submanifold



Riemannian Optimization

Roughly, a Riemannian manifold $\mathcal{M}$ is a smooth set with a smoothly-varying inner product on the tangent spaces.

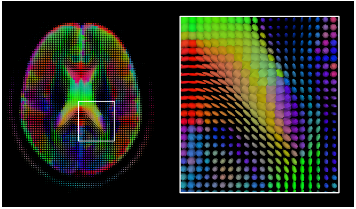


Riemannian manifold = Manifold + Riemannian metric (inner products)

Applications

Embedded submanifold: Computation on SPD manifold

- SPD manifold: $\mathcal{S}^n_{++} = \{X \in \mathbb{R}^{n\times n} : X = X^T, X \succ 0\}$;
- Applications of SPD matrices
  - Diffusion tensors in medical imaging [CSV12, FJ07, RTM07]
  - Describing images and video [LWM13, SFD02, ASF$^+$05, TPM06, HWSC15]



- Motivation of averaging SPD matrices
  - denoising / interpolation
  - clustering / classification

Applications

Embedded submanifold: Computation on SPD manifold

One averaging SPD matrices method:

$$G(A_1, \dots, A_k) = \arg \min_{X \in \mathcal{S}_{++}^n} \frac{1}{2k} \sum_{i=1}^{k} \mathrm{dist}^2(X, A_i),$$

where $\mathrm{dist}(X, Y) = \| \log(X^{-1/2} Y X^{-1/2}) \|_F$ is the distance under the Riemannian metric $\langle \eta_X, \xi_X \rangle = \mathrm{trace}(\eta_X X^{-1} \xi_X X^{-1})$.

Applications

Embedded submanifold: Computation on SPD manifold

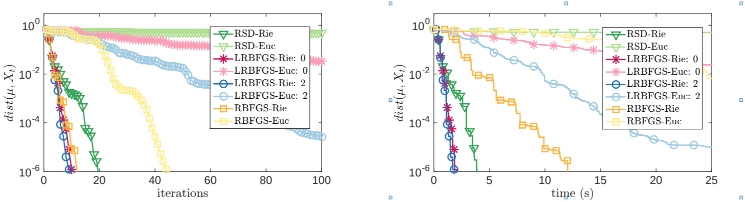
One averaging SPD matrices method:

$$G(A_1, \dots, A_k) = \arg \min_{X \in \mathcal{S}_{++}^n} \frac{1}{2k} \sum_{i=1}^{k} \mathrm{dist}^2(X, A_i),$$

where $\mathrm{dist}(X, Y) = \|\log(X^{-1/2} Y X^{-1/2})\|_F$ is the distance under the Riemannian metric $\langle \eta_X, \xi_X \rangle = \mathrm{trace}(\eta_X X^{-1} \xi_X X^{-1})$.

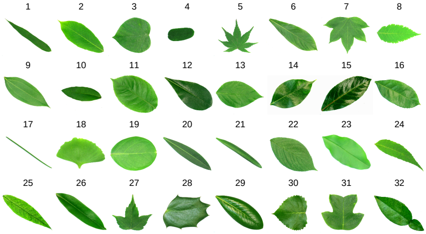
Why shall we use Riemannian optimization approach?

- $K = 30$, $n = 100$, and $1 \leq \kappa(A_i) \leq 10^5$.



Applications

Quotient manifold: Computation on shape space



- Classification [LKS+12, HGSA15]
- Face recognition [DBS+13]

Applications

Quotient manifold: Computation on shape space

- Elastic shape analysis invariants:
  - Rescaling
  - Translation
  - Rotation
  - Reparametrization
- The shape space is a quotient space

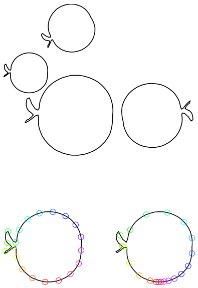
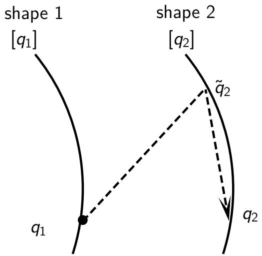


Figure: All are the same shape.

Applications

Quotient manifold: Computation on shape space

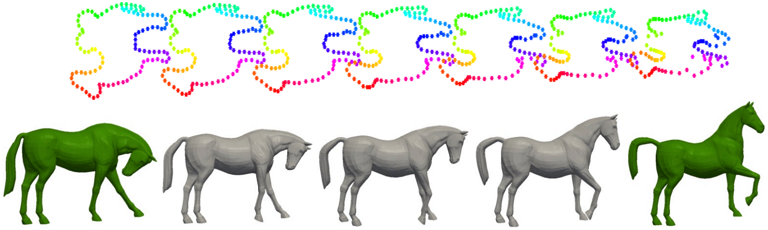
Registration



- Optimization problem $\min_{q_2 \in [q_2]} \mathrm{dist}(q_1, q_2)$ is defined on a Riemannian manifold

Applications

Quotient manifold: Computation on shape space
Geodesic / Interpolation



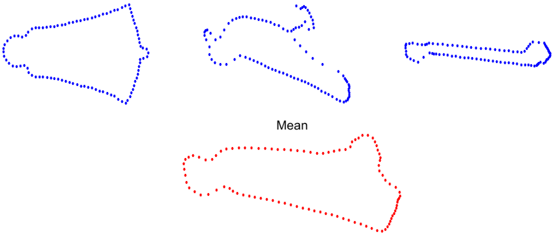
$$\min_{\alpha \in \mathcal{H}_{x,y}} \frac{1}{2} \int_0^1 \langle \dot{\alpha}(\tau), \dot{\alpha}(\tau) \rangle \mathrm{d}\tau$$

- Computation of a geodesic between two shapes
- Interpolation in shape space

Applications

Quotient manifold: Computation on shape space

Karcher mean



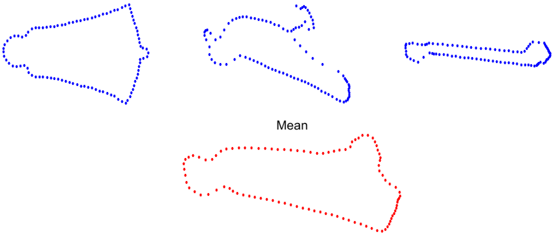
$$\min_{X \text{ is a shape}} \frac{1}{2k} \sum_{i=1}^{k} \mathrm{dist}^2(X, S_i),$$

- Computation of Karcher mean of a population of shapes

Applications

Quotient manifold: Computation on shape space

Karcher mean



$$\min_{X \text{ is a shape}} \frac{1}{2k} \sum_{i=1}^{k} \mathrm{dist}^2(X, S_i),$$

- Computation of Karcher mean of a population of shapes

Riemannian optimization is used since these problems naturally involve a Riemannian manifold

Smooth Optimization Framework

Iterations on the Manifold

Consider the following generic update for an iterative Euclidean optimization algorithm:

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

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

- Steepest descent: $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$
- Newton's method: $x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$
- Trust region method: Δx_k is set by optimizing a local model.

x_k $x_k + d_k$

Riemannian Manifolds Provide

- Riemannian concepts describing directions and movement on the manifold
- Riemannian analogues for gradient and Hessian

Smooth Optimization Framework

Riemannian gradient and Riemannian Hessian

Definition

The Riemannian gradient of f at x is the unique tangent vector in $\mathrm{T}_x\mathcal{M}$ satisfying $\forall \eta \in \mathrm{T}_x\mathcal{M}$, the directional derivative

$$\mathrm{D}f(x)[\eta] = g(\mathrm{grad} f(x), \eta)$$

and $\mathrm{grad} f(x)$ is the direction of steepest ascent.

Definition

The Riemannian Hessian of f at x is a symmetric linear operator from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_x\mathcal{M}$ defined as

$$\mathrm{Hess} f(x) : \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M} : \eta \to \nabla_\eta \mathrm{grad} f,$$

where ∇ is the affine connection.

Smooth Optimization Framework

Retractions

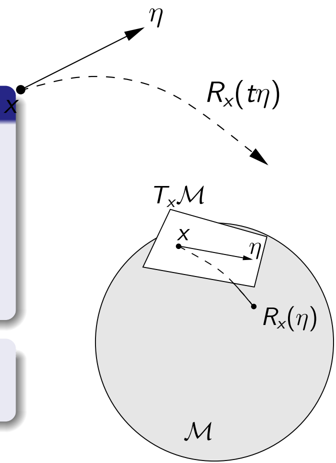
Euclidean	Riemannian
$x_{k+1} = x_k + \alpha_k d_k$	$x_{k+1} = R_{x_k}(\alpha_k \eta_k)$

Definition

A retraction is a mapping R from $\mathrm{T}\mathcal{M}$ to $\mathcal{M}$ satisfying the following:

- R is continuously differentiable
- $R_x(0) = x$
- $\mathrm{D}R_x(0)[\eta] = \eta$

- maps tangent vectors back to the manifold
- defines curves in a direction



Smooth Optimization Framework

Categories of Riemannian smooth optimization methods

Retraction-based: local information only

Line search-based: use local tangent vector and $R_x(t\eta)$ to define line

- Steepest decent
- Newton

Local model-based: series of flat space problems

- Riemannian trust region Newton (RTR)
- Riemannian adaptive cubic overestimation (RACO)

Smooth Optimization Framework

Categories of Riemannian smooth optimization methods

Retraction and transport-based: information from multiple tangent spaces

- Nonlinear conjugate gradient: multiple tangent vectors
- Quasi-Newton e.g. Riemannian BFGS: transport operators between tangent spaces

Additional element required for optimizing a cost function;

- formulas for combining information from multiple tangent spaces.

Smooth Optimization Framework

Categories of Riemannian smooth optimization methods

Retraction and transport-based: information from multiple tangent spaces

- Nonlinear conjugate gradient: multiple tangent vectors
- Quasi-Newton e.g. Riemannian BFGS: transport operators between tangent spaces

Additional element required for optimizing a cost function;

- formulas for combining information from multiple tangent spaces.

Vector Transport:

- Vector transport: Transport a tangent vector from one tangent space to another;
- $\mathcal{T}_{\eta_x}\xi_x$, denotes transport of ξ_x to tangent space of $R_x(\eta_x)$. R is a retraction associated with $\mathcal{T}$;

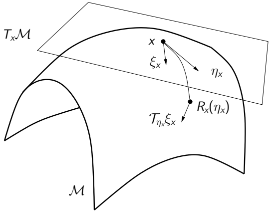


Figure: Vector transport.

Smooth Optimization Framework

Retraction/Transport-based Riemannian Optimization

Given a retraction and a vector transport, we can generalize classical unconstrained smooth optimization methods from Euclidean space to the Riemannian manifold.

Smooth Optimization Framework

Retraction/Transport-based Riemannian Optimization

Given a retraction and a vector transport, we can generalize classical unconstrained smooth optimization methods from Euclidean space to the Riemannian manifold.

Do the Riemannian versions of those methods work well?

Smooth Optimization Framework

Retraction/Transport-based Riemannian Optimization

Given a retraction and a vector transport, we can generalize classical unconstrained smooth optimization methods from Euclidean space to the Riemannian manifold.

Do the Riemannian versions of those methods work well?

No, generally

- Lose many theoretical results and important properties;
- Impose restrictions on retraction/vector transport;

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Manifold recognition
- Riemannian metric
- Retraction / Geodesic
- Vector transport / Parallel translation

[EAS1998] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, 1998

[CMV2017] T Carson, D. G. Mixon, and S. Villar. Manifold optimization for k-means clustering. In *2017 International Conference on Sampling Theory and Applications (SampTA)*, 73–77. IEEE, 2017

[SDN2021] G. Song, W. Ding, and M. K. Ng, Low rank pure quaternion approximation for pure quaternion matrices, *SIAM Journal on Matrix Analysis and Applications*, 42, pp. 58–82, 2021

[VAV2013] B. Vandereycken, P.-A. Absil, and S. Vandewalle. A Riemannian geometry with complete geodesics for the set of positive semidefinite matrices of fixed rank, *IMA Journal of Numerical Analysis*, 33.2, 481–514, 2013.

[Zim2017] R. Zimmermann. A matrix-algebraic algorithm for the Riemannian logarithm on the Stiefel manifold under the canonical metric. *SIAM Journal on Matrix Analysis and Applications*, 38.2, 322–342, 2017.

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Smooth unconstrained optimization algorithms
- Nonsmooth unconstrained optimization algorithms
- Constrained optimization algorithms

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Smooth unconstrained optimization algorithms
- Nonsmooth unconstrained optimization algorithms
- Constrained optimization algorithms

Riemannian optimization mainly focuses on this topic.
Discuss later.

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Computations on the SPD manifold;
- Computations on the shape space;
- Clustering and graph partitions;
- Beamforming in wireless communication;
- Blind source separation;
- etc

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Representation of a manifold and tangent spaces;
- Choose a Riemannian metric;
- Choose a retraction;
- Choose a vector transport;

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Representation of a manifold and tangent spaces;
- Choose a Riemannian metric;
- Choose a retraction;
- Choose a vector transport;

Above factors may influence algorithms significantly.

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

- Manopt (Matlab library) [Boumal, Mishra, Absil, Sepulchre(2014)]
- Pymanopt (Python version of Manopt) [Townsend, Koep, Weichwald (2016)]
- Manoptjl (Julia, nonsmooth methods) [Bergmann (2019)]
- ROPTLIB (C++ library, interfaces to Matlab and Julia) [Huang, Absil, Gallivan, Hand (2018)]
- ManifoldOptim (R wrapper of ROPTLIB) [Martin, Raim, Huang, Adragni (2018)]
- McTorch (Python, GPU acceleration) [Meghawanshi, Jawanpuria, Kunchukuttan, Kasai, Mishra (2018)]
- CDOpt (Python, embedded submanifold in the form of $c(x) = 0$) [Xiao, Hu, Liu, Toh (2022)]

Research Foci of Riemannian Optimization

1. Manifold recognition, geometry structure analyses and computations;
2. Generalization Euclidean algorithms to the Riemannian setting;
3. Algorithms specialization for applications;
4. Library developments;

Provide theories to explain behaviors of existing algorithms for particular applications

- [MBDG2023]: IRKA is a Riemannian gradient descent method;
- [YHAG2020]: Richardson-like iteration for matrix geometric mean is a Riemannian gradient descent method;
- [BM2006]: The improved BFGS method is a Riemannian BFGS method using vector transport by parallelization;

[MBDG2023] P. Mlinaric, C. Beattie, Z. Drmac, and S. Gugercin. IRKA is a Riemannian Gradient Descent Method. arxiv:2311.02031, 2023

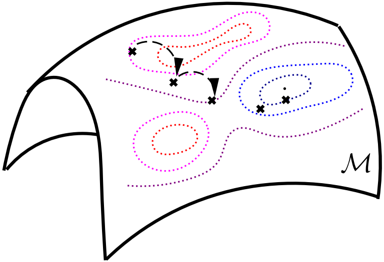
[YHAG2020] X. Yuan, W. Huang, P.-A. Absil, K. A. Gallivan. Computing the matrix geometric mean: Riemannian vs Euclidean conditioning, implementation techniques, and a Riemannian BFGS method, *Numerical Linear Algebra with Applications*, 27:5, 1-23, 2020

[BM2006] I. Brace and J. H. Manton. An improved BFGS-on-manifold algorithm for computing weighted low rank approximations. *Proceedings of 17th international Symposium on Mathematical Theory of Networks and Systems*, P.1735–1738, 2006

Comparison with Constrained Optimization

Not all Riemannian optimization problem can be formulated as constrained optimization problems, and vice versa.

- All iterates on the manifold
- Convergence properties of unconstrained optimization algorithms
- No need to consider Lagrange multipliers or penalty functions
- Exploit the structure of the constrained set



A Non-exhaustive Review

- Smooth unconstrained problems
  - Steepest descent: Smith 1994; Helmke-Moore 1994; Iannazzo-Porcelli 2019;
  - Conjugate gradient: Smith 1994; Gallivan-Absil 2010; Ring-Wirth 2012; Sato-Iwai 2015;
  - Quasi-Newton: Ring-Wirth 2012; Huang-Absil-Gallivan 2018; Huang-Gallivan 2022
  - Newton-CG: Absil-Baker-Gallivan 2007; Huang-Huang 2023
- Nonsmooth unconstrained problems
  - Proximal point method: Ferreira-Oliveira 2002;
  - Optimality conditions: Yang-Zhang-Song 2014;
  - Gradient sampling: Huang 2013; Hosseini and Uschmajew 2017;
  - ϵ-subgradient-based methods: Grohs-Hosseini 2015;
  - Proximal gradient methods: Huang-Wei 2022;
  - Proximal Newton method: Si-Absil-Huang-Jiang-Vary 2023;
- Constrained problems:
  - Augmented Lagrangian methods: Boumal-Liu 2019;
  - Sequential quadratic programming: Obara-Okuno-Takeda 2022;
  - Frank-Wolfe Methods: Weber-Sra 2023;

A Non-exhaustive Review

- Smooth unconstrained problems:
  - Stiefel manifold: Wen-Yin 2012; Jiang-Dai 2014; Xiao-Liu-Yuan 2020; Dai-Wang-Zhou 2020
  - Symplectic Stiefel manifold: Gao-Son-Absil-Stykel 2021
  - Symmetric positive definite manifold: Bini-Iannazzo 2013; Zhang 2017; Yuan-Huang-Absil-Gallivan 2020;
  - Fixed rank manifold: Wen-Yin-Zhang 2012; Mishra 2014; Sutti-Vandereycken 2021; Levin-Kileel-Boumal 2022
- Nonsmooth unconstrained problems:
  - Stiefel Manifold: Huang-Wei 2019; Chen-Ma-So-Zhang 2020; Xiao-Liu-Yuan 2020;
  - Fixed rank manifold: Cambier-Absil 2016;
  - Matrix manifolds: Zhou-Bao-Ding-Zhu 2022
  - Smooth equation constraints: Xiao-Liu-Toh 2023
- Constrained problems:
  - Stiefel + non-negativity: Jiang-Meng-Wen-Chen 2019;
  - Symmetric positive definite + zeros: Phan-Menickelly 2020;

Content

Main references

- *An Introduction to Optimization on Smooth Manifolds*, Nicolas Boumal, Version of May 25 2020

  A nice introduction for readers who do not have background on Riemannain manifold

- *Optimization Algorithms on Matrix Manifolds*, P.-A. Absil, R. Mahony, R. Sepulchre, Princeton University Press, January 2008

  More suitable for readers who are familiar with Riemannian manifold.

Main references

- *An Introduction to Optimization on Smooth Manifolds*, Nicolas Boumal, Version of May 25 2020

  A nice introduction for readers who do not have background on Riemannain manifold $\Longleftarrow$

- *Optimization Algorithms on Matrix Manifolds*, P.-A. Absil, R. Mahony, R. Sepulchre, Princeton University Press, January 2008

  More suitable for readers who are familiar with Riemannian manifold.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space

- Linear spaces: $\mathbb{R}^n$, $\mathbb{R}^{n \times p}$, $\mathrm{Sym}(n)$, $\mathbb{C}^n$, $\mathbb{R}^{n_1 \times n_2 \times n_3}$
- Euclidean space: A linear space equipped with a Euclidean metric

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds

Definition

Let $\mathcal{M}$ be a subset of a linear space $\mathcal{E}$. We say $\mathcal{M}$ is a embedded submanifold of $\mathcal{E}$ if either of the following holds:

1. $\mathcal{M}$ is an open subset of $\mathcal{E}$. Then we also call $\mathcal{M}$ an open submanifold. if $\mathcal{M} = \mathcal{E}$, we also call it a linear manifold.
2. For a fixed integer $k \geq 1$ and for each $x \in \mathcal{M}$ there exists a neighborhood $\mathcal{U}$ of x in $\mathcal{E}$ and a smooth function $h : \mathcal{U} \to \mathbb{R}^k$ such that
   - If y is in $\mathcal{U}$, then $h(y) = 0$ if and only if $y \in \mathcal{M}$; and
   - $\mathrm{rank} \mathrm{D}h(x) = k$.

   Such a function h is called a local defining function for $\mathcal{M}$ at x.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds

Example

Show that the unique sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : x^T x = 1\}$ is an embedded submanifold.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds

Example

Show that $\{(x_1, x_2) : x_1^2 - x_2^2 = 0\}$ is not an embedded submanifold.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds
- Topology

Definition

A subset $\mathcal{U}$ of $\mathcal{M}$ is open in $\mathcal{M}$ if $\mathcal{U}$ is the intersection of $\mathcal{M}$ with an open subset of $\mathcal{E}$. This is called the subspace topology.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds
- Topology

Proposition

Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Any open subset of $\mathcal{M}$ is also an embedded (but not necessarily open) submanifold of $\mathcal{E}$, with same dimension and tangent spaces as $\mathcal{M}$.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds
- Topology

Proposition

Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Any open subset of $\mathcal{M}$ is also an embedded (but not necessarily open) submanifold of $\mathcal{E}$, with same dimension and tangent spaces as $\mathcal{M}$.

Example

Show that $\Delta_+^{n+1} = \{x \in \mathbb{R}^n : x_1 + \cdot + x_n = 1 \text{ and } x_1, \ldots, x_n > 0\}$ is an embedded submanifold.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds
- Topology
- Tangent space

Definition

Let $\mathcal{M}$ be a subset of $\mathcal{E}$. For all $x \in \mathcal{M}$, define:

$$\mathrm{T}_x\mathcal{M} = \{c'(0) : c : I \to \mathcal{M} \text{ is smooth around 0 and } c(0) = x\},$$

where I is an open subset of $\mathbb{R}$.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds
- Topology
- Tangent space

Theorem

Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Consider $x \in \mathcal{M}$ and the set $\mathrm{T}_x\mathcal{M}$. If $\mathcal{M}$ is an open submanifold, then $\mathrm{T}_x\mathcal{M} = \mathcal{E}$. Otherwise, $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$ with h any local defining function at x.

Embedded submanifolds

Definitions

What is a smooth Riemannian manifold? (focus on Embedded submanifolds)

- Embedding space: Euclidean space
- Embedded submanifolds
- Topology
- Tangent space

Example

Derive an expression of $\mathrm{T}_x\mathbb{S}^{n-1}$, where $\mathbb{S}^{n-1}$ is the unique sphere.

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions

Definition

Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$. A map $F : \mathcal{M} \to \mathcal{M}'$ is smooth if and only if it admits a smooth extension $\overline{F} : \mathcal{U} \to \mathcal{E}'$ in a neighborhood $\mathcal{U}$ of $\mathcal{M}$ in $\mathcal{E}$, so that $F(x) = \overline{F}(x)$ for all $x \in \mathcal{M}$, that is, $\overline{F}$ is the restriction of $\overline{F}$ to $\mathcal{M}$: $F = \overline{F}|_{\mathcal{M}}$

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions

Example

- Given a smooth extension of the function $f\colon \mathbb{S}^{n-1} \to \mathbb{R} : x \mapsto x^T A x$ on the unique sphere.
- Given an example of an embedded submanifold $\mathcal{M}$ and a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ for which where does not exist a smooth extension $\bar{f}\colon \mathcal{E} \to \mathbb{R}$ smooth on all of $\mathcal{E}$.

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions
- Differential

Definition

The differential of $F : \mathcal{M} \to \mathcal{M}'$ at x is a linear operator $\mathrm{D}F(x) : \mathrm{T}_x\mathcal{M} \to \mathrm{T}_{F(x)}\mathcal{M}'$ defined by:

$$\mathrm{D}F(x)[v] = \frac{d}{dt} F(c(t))|_{t=0},$$

where c is a smooth curve on $\mathcal{M}$ passing through x at $t = 0$ with velocity v.

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions
- Differential

Proposition

Let $\overline{F}$ be a smooth extension of $F : \mathcal{M} \to \mathcal{M}'$. Therefore,
$\mathrm{D}F(x) = \mathrm{D}\overline{F}(x)|_{\mathrm{T}_x\mathcal{M}}$.

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions
- Differential

Proposition

Let $\overline{F}$ be a smooth extension of $F : \mathcal{M} \to \mathcal{M}'$. Therefore, $\mathrm{D}F(x) = \mathrm{D}\overline{F}(x)|_{\mathrm{T}_x\mathcal{M}}$.

- $DF(x)$ is linear since $D\overline{F}(x)$ is linear.
- The definition of $DF(x)$ is independent of c with $c(0) = x$ and $c'(0) = v$.

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions
- Differential

Example

Let $f : \mathbb{S}^{n-1} \to \mathbb{R} : x \mapsto x^T A x$, where $A = A^T$. Compute $\mathrm{D}f(x)[v]$ for $v \in \mathrm{T}_x\mathbb{S}^{n-1}$.

Embedded submanifolds

Smooth maps

Smooth maps on smooth manifolds

- Definitions
- Differential
- Properties of differentials

Example

Let $f : \mathcal{M} \to \mathbb{R}$, $F : \mathcal{M} \to \mathcal{M}'$ and $G : \mathcal{M}' \to \mathcal{M}''$ be smooth, where $\mathcal{M}$, $\mathcal{M}'$, and $\mathcal{M}''$ are embedded submanifolds of $\mathcal{E}$, $\mathcal{E}'$ and $\mathcal{E}''$ respectively. Then

- Show that $fF : x \to f(x)F(x)$ is smooth from $\mathcal{M}$ to $\mathcal{E}'$ and we have product rule:

$$D(fF)(x)[v] = F(x)\mathrm{D}f(x)[v] + f(x)\mathrm{D}F(x)[v].$$

- Show that $\mathrm{D}(G \circ F)(x)[v] = \mathrm{D}G(F(x))[\mathrm{D}F(x)[v]]$.

Embedded submanifolds

Tangent bundle and vector fields

Tangent bundle and vector fields

Embedded submanifolds

Tangent bundle and vector fields

Tangent bundle and vector fields

- Tangent bundle: Definition

Definition

The tangent bundle of a manifold $\mathcal{M}$ is the disjoint union of the tangent spaces of $\mathcal{M}$:

$$\mathrm{T}\mathcal{M} = \{(x, v) : x \in \mathcal{M} \text{ and } v \in \mathrm{T}_x\mathcal{M}\}.$$

Note that the first component x is used for the notion of "disjoint" union when $\mathcal{M}$ is en embedded submanifold. In the abstract definition of the tangent bundle, tangent spaces are disjoint by definition. It follows that $\mathrm{T}\mathcal{M} = \cup_{x\in\mathcal{M}}\mathrm{T}_x\mathcal{M}$.

Embedded submanifolds

Tangent bundle and vector fields

Tangent bundle and vector fields

- Tangent bundle: Definition

Theorem

If $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$, the tangent bundle $\mathrm{T}\mathcal{M}$ is an embedded submanifold of $\mathcal{E} \times \mathcal{E}$ of dimension $2\dim(\mathcal{M})$.

Embedded submanifolds

Tangent bundle and vector fields

Tangent bundle and vector fields

- Tangent bundle: Definition
- Vector fields: Definition

Definition

A vector field on a manifold $\mathcal{M}$ is a map $V : \mathcal{M} \to \mathrm{T}\mathcal{M}$ such that $V(x)$ is in $\mathrm{T}_x\mathcal{M}$ for all $x \in \mathcal{M}$. If V is a smooth map, we say it is a smooth vector field. The set of smooth vector fields is denoted by $\mathfrak{X}(\mathcal{M})$.

Embedded submanifolds

Tangent bundle and vector fields

Tangent bundle and vector fields

- Tangent bundle: Definition
- Vector fields: Definition

Proposition

For $\mathcal{M}$ an embedded submanifold of $\mathcal{E}$, a vector field V on $\mathcal{M}$ is smooth if and only if there exists a smooth vector field $\overline{V}$ on $\mathcal{U} \subseteq \mathcal{E}$ (a neighborhood of $\mathcal{M}$) such that $V = \overline{V}|_{\mathcal{M}}$.

Embedded submanifolds

Tangent bundle and vector fields

Tangent bundle and vector fields

- Tangent bundle: Definition
- Vector fields: Definition

Example

Give a smooth vector field on $\mathbb{S}^{n-1}$, where $\mathbb{S}^{n-1}$ is the unit sphere.

Embedded submanifolds

Retraction

Retraction

Embedded submanifolds

Retraction

Retraction

- Definition

A retraction on $\mathcal{M}$ is a smooth map $R : \mathrm{T}\mathcal{M} \to \mathcal{M}$ with the following properties. For each $x \in \mathcal{M}$, let $R_x : \mathrm{T}_x\mathcal{M} \to \mathcal{M}$ be the restriction of R at x, so that $R_x(v) = R(x, v)$. Then

1. $R_x(0) = x$, and
2. $\mathrm{D}R_x(0) : \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ is the identity map: $\mathrm{D}R_x(0)[v] = v$.

Equivalently, each curve $c(t) = Rx(tv)$ satisfies $c(0) = x$ and $c'(0) = v$.

Embedded submanifolds

Retraction

Retraction

- Definition

Example

Show that below two maps are retractions on $\mathbb{S}^{n-1}$

- $R_x(v) = \frac{x+v}{\|x+v\|}$, and
- $R_x(v) = \cos(\|v\|)x + \sin(\|v\|)\frac{v}{\|v\|}$,

where $v \in \mathrm{T}_x\mathbb{S}^{n-1}$.

Embedded submanifolds

Retraction

Retraction

- Definition
- Recognition

Let $\mathcal{M}$ be an embedded manifold of a vector space $\mathcal{E}$ and let $\mathcal{N}$ be an abstract manifold such that $\dim(\mathcal{M}) + \dim(\mathcal{N}) = \dim(\mathcal{E})$. Assume that there is a diffeomorphism $\phi : \mathcal{M} \times \mathcal{N} \to \mathcal{E}_* : (F, G) \mapsto \phi(F, G)$, where $\mathcal{E}_*$ is an open subset of $\mathcal{E}$, with a neural element I in $\mathcal{N}$ satisfying $\phi(F, I) = F, \forall F \in \mathcal{M}$.

Theorem [AMS08, Proposition 4.1.2]

The mapping

$$R_X(\xi) = \pi_1(\phi^{-1}(X + \xi))$$

defines a retraction $\mathcal{M}$, where $\pi_1 : \mathcal{M} \times \mathcal{N} \to \mathcal{M} : (F, G) \mapsto F$ is the projection onto the first component.

Embedded submanifolds

Retraction

Retraction

- Definition
- Recognition
- Construction

- Exponential mapping: $R_x(v) = \gamma(1)$, where $\gamma(t)$ is the geodesic such that $\gamma(0) = x$ and $\gamma'(0) = v$.
- Retraction by projection: $R_x(v) = \arg\min_{y \in \mathcal{M}} \|x + v - y\|$
- Orthographic retraction: $R_x(v) = x + v + u$, where $u \in \mathrm{T}_x^\perp \mathcal{M}$.

See [AM12].

Embedded submanifolds

Riemannian metric

Riemannian metric

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition

Definition

An inner product on $\mathrm{T}_x\mathcal{M}$ is a bilinear, symmetric, positive definite function $\langle\cdot,\cdot\rangle : \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathbb{R}$. It induces a norm for tangent vectors: $\|u\|_x = \sqrt{\langle u, u\rangle}$. A metric on $\mathcal{M}$ is a choice of inner product $\langle\cdot,\cdot\rangle$ for each $x \in \mathcal{M}$.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition

Definition

An inner product on $\mathrm{T}_x\mathcal{M}$ is a bilinear, symmetric, positive definite function $\langle\cdot,\cdot\rangle : \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathbb{R}$. It induces a norm for tangent vectors: $\|u\|_x = \sqrt{\langle u, u\rangle}$. A metric on $\mathcal{M}$ is a choice of inner product $\langle\cdot,\cdot\rangle$ for each $x \in \mathcal{M}$.

Definition

A metric $\langle\cdot,\cdot\rangle$ on $\mathcal{M}$ is a Riemannian metric if it varies smoothly with x in the sense that if V, W are two smooth vector fields on $\mathcal{M}$ then the function $x \mapsto \langle V(x), W(x)\rangle$ is smooth from $\mathcal{M}$ to $\mathbb{R}$.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold

Definition

A manifold with a Riemannian metric is a Riemannian manifold.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold

Proposition

Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$, and let $\langle \cdot, \cdot \rangle$ be the Euclidean metric on $\mathcal{E}$. Then, the metric on $\mathcal{M}$ defined at each x by restriction, $\langle u, v \rangle = \langle u, v \rangle$ for $u, v \in \mathrm{T}_x\mathcal{M}$ is a Riemannian metric.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold

Proposition

Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$, and let $\langle \cdot, \cdot \rangle$ be the Euclidean metric on $\mathcal{E}$. Then, the metric on $\mathcal{M}$ defined at each x by restriction, $\langle u, v \rangle = \langle u, v \rangle$ for $u, v \in \mathrm{T}_x\mathcal{M}$ is a Riemannian metric.

Definition

Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$. Equipped with the Riemannian metric obtained by restriction of the metric of $\mathcal{E}$, we call $\mathcal{M}$ a Riemannian submanifold of $\mathcal{E}$.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold
- Riemannian gradient

Definition

Let $f\colon \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold $\mathcal{M}$. The Riemannian gradient of f is the vector field $\mathrm{grad} f$ on $\mathcal{M}$ uniquely defined by these identities:

$$\forall (x, v) \in \mathrm{T}\mathcal{M}, \mathrm{D}f(x)[v] = \langle v, \mathrm{grad} f(x) \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the Riemannian metric.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold
- Riemannian gradient

Proposition

Let $\mathcal{M}$ be a Riemannian submanifold of $\mathcal{E}$ endowed with the Euclidean metric $\langle \cdot, \cdot \rangle$ and let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function. The Riemannian gradient of f is given by

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad} \bar{f}(x)),$$

where $\bar{f}$ is a smooth extension of f to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$, and Proj_x denotes the orthogonal projection to $\mathrm{T}_x\mathcal{M}$.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold
- Riemannian gradient

Example

Derive the Riemannnian gradient of $f\colon \mathbb{S}^{n-1} \to \mathbb{R} : x \mapsto x^T A x$, where the Riemannian metric is the Euclidean metric and $A = A^T$.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold
- Riemannian gradient

Example

Show that the Riemannian gradient is unique by its definition.

Embedded submanifolds

Riemannian metric

Riemannian metric

- Definition
- Riemannian manifold, Riemannian submanifold
- Riemannian gradient

Example

Consider the relative interior of the simplex,

$$\mathcal{M} = \Delta_+^{n-1} = \{x \in \mathbb{R}^n : x_1, \ldots, x_n > 0 \text{ and } x_1 + \cdots + x_n = 1\},$$

as an embedded submanifold of $\mathbb{R}^n$. Its tangent spaces are given by

$$\mathrm{T}_x\mathcal{M} = \{v \in \mathbb{R}^n : v_1 + \cdots + v_n = 0\}.$$

Show that $\langle u, v\rangle = \sum_{i=1}^n \frac{u_i v_i}{x_i}$ defines a Riemannian metric on $\mathcal{M}$. Then, considering a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ and a smooth extension $\bar{f}$ on a neighborhood of $\mathcal{M}$ in $\mathbb{R}^n$ (equipped with the Euclidean metric), give an expression for $\mathrm{grad} f(x)$ in terms of $\mathrm{grad} \bar{f}(x)$.

Embedded submanifolds

A Riemannian steepest descent algorithm

A Riemannian steepest descent algorithm

Embedded submanifolds

A Riemannian steepest descent algorithm

A Riemannian steepest descent algorithm

- A Riemannian gradient descent algorithm

A representative Riemannian steepest descent algorithm:

1: Given: Initial step size $\sigma > 0$; Shrinking parameter $\rho \in (0, 1)$; $c \in (0, 1)$; initial iterate $x_0 \in \mathcal{M}$;
2: **for** k = 0, 1, 2,... **do**
3: Find the largest $\alpha \in \{\sigma, \sigma\rho, \sigma\rho^2, \dots, \}$ such that

$$f(R_{x_k}(-\alpha\,\mathrm{grad} f(x_k))) \leq f(x_k) - c\alpha \|\mathrm{grad} f(x_k)\|_{x_k}^2$$

4: Set $x_{k+1} = R_{x_k}(-\alpha\,\mathrm{grad} f(x_k))$.
5: **end for**

Embedded submanifolds

A Riemannian steepest descent algorithm

A Riemannian steepest descent algorithm

- A Riemannian gradient descent algorithm
- First-order optimality conditions

Theorem

Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold. If x is a local minimizer of f, then $\mathrm{grad} f(x) = 0$.

Embedded submanifolds

A Riemannian steepest descent algorithm

A Riemannian steepest descent algorithm

- A Riemannian gradient descent algorithm
- First-order optimality conditions

Theorem

Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold. If x is a local minimizer of f, then $\mathrm{grad} f(x) = 0$.

Definition

Given a smooth function f on a Riemannian manifold $\mathcal{M}$, we call $x \in \mathcal{M}$ a critical point or a stationary point of f is $\mathrm{grad} f(x) = 0$.

Embedded submanifolds

A Riemannian steepest descent algorithm

A Riemannian steepest descent algorithm

- A Riemannian gradient descent algorithm
- First-order optimality conditions
- Convergence

Assumption

1. *There exists* $f_{\mathrm{low}} \in \mathbb{R}$ *such that* $f(x) \geq f_{\mathrm{low}}$ *for all* $x \in \mathcal{M}$.
2. *For a given subset* S *of the tangent bundle* $\mathrm{T}\mathcal{M}$, *there exists a constant* $L > 0$ *such that, for all* $(x, s) \in S$,

$$f(R_x(s)) \leq f(x) + \langle \mathrm{grad} f(x), s \rangle + \frac{L}{2} \|s\|^2.$$

Embedded submanifolds

A Riemannian steepest descent algorithm

A Riemannian steepest descent algorithm

- A Riemannian gradient descent algorithm
- First-order optimality conditions
- Convergence

Theorem

Let f be a smooth function satisfying the assumptions. Let $\{(x_i, s_i)\}_{i=0}^{\infty}$ be the pairs generated by the algorithm. If these pairs are in S defined in the assumption, then

$$\lim_{k\to\infty} \|\mathrm{grad} f(x_k)\| = 0.$$

References I

P.-A. Absil and J. Malick.
Projection-like retractions on matrix manifolds.
SIAM Journal on Optimization, 22(1):135–158, 2012.

P.-A. Absil, R. Mahony, and R. Sepulchre.
Optimization algorithms on matrix manifolds.
Princeton University Press, Princeton, NJ, 2008.

Ognjen Arandjelovic, Gregory Shakhnarovich, John Fisher, Roberto Cipolla, and Trevor Darrell.
Face recognition with image sets using manifold density divergence.
In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 581–588. IEEE, 2005.

Guang Cheng, Hesamoddin Salehian, and Baba Vemuri.
Efficient recursive algorithms for computing the mean diffusion tensor and applications to DTI segmentation.
Computer Vision–ECCV 2012, pages 390–401, 2012.

H. Drira, B. Ben Amor, A. Srivastava, M. Daoudi, and R. Slama.
3D face recognition under expressions, occlusions, and pose variations.
Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35(9):2270–2283, 2013.

P. T. Fletcher and S. Joshi.
Riemannian geometry for the statistical analysis of diffusion tensor data.
Signal Processing, 87(2):250–262, 2007.

W. Huang, K. A. Gallivan, Anuj Srivastava, and P.-A. Absil.
Riemannian optimization for registration of curves in elastic shape analysis.
Journal of Mathematical Imaging and Vision, 54(3):320–343, 2015.
DOI:10.1007/s10851-015-0606-8.

References II

Zhiwu Huang, Ruiping Wang, Shiguang Shan, and Xilin Chen.
Face recognition on large-scale video in the wild with hybrid Euclidean-and-Riemannian metric learning.
Pattern Recognition, 48(10):3113–3124, 2015.

H. Laga, S. Kurtek, A. Srivastava, M. Golzarian, and S. J. Miklavcic.
A Riemannian elastic metric for shape-based plant leaf classification.
2012 International Conference on Digital Image Computing Techniques and Applications (DICTA), pages 1–7, December 2012. doi:10.1109/DICTA.2012.6411702.

Jiwen Lu, Gang Wang, and Pierre Moulin.
Image set classification using holistic multiple order statistics features and localized multi-kernel metric learning.
In *Proceedings of the IEEE International Conference on Computer Vision*, pages 329–336, 2013.

Y. Rathi, A. Tannenbaum, and O. Michailovich.
Segmenting images on the tensor manifold.
In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.

Gregory Shakhnarovich, John W Fisher, and Trevor Darrell.
Face recognition from long-term observations.
In *European Conference on Computer Vision*, pages 851–865. Springer, 2002.

Oncel Tuzel, Fatih Porikli, and Peter Meer.
Region covariance: A fast descriptor for detection and classification.
In *European conference on computer vision*, pages 589–600. Springer, 2006.