Hajg Jasa.

- Studied both Physics and Mathematics at University of Milano-Bicocca
- main focus: Differential Geometry
- Reserach Internship at University of Zurich: statistical modelling and data analysis
- Since August 2022: PhD student at IMF

Thesis Project.
Nonsmooth Optimization on Riemannian manifolds

Goal, Find (numerically / fast)

$$\operatorname*{arg\,min}_{p \in \mathcal{M}} f(p),$$

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

Goal, Find (numerically / fast)

$$\operatorname*{arg\,min}_{p\in\mathcal{M}} f(p),$$

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

Challenges. Adapt algorithms, since $\mathcal{M}$ is nonlinear.

- Find is a "good replacement" $+$
- no "global space of directions"; $T_p\mathcal{M} \neq T_q\mathcal{M}$

NTNU

Goal, Find (numerically / fast)

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

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

Challenges. Adapt algorithms, since $\mathcal{M}$ is nonlinear.

- Find is a “good replacement” $+$
- no “global space of directions”; $T_p\mathcal{M} \neq T_q\mathcal{M}$

Task. How to handle the case when f is nonsmooth?

NTNU

Goal, Find (numerically / fast)

$$\operatorname*{arg\,min}_{p\in\mathcal{M}} f(p),$$

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

Challenges. Adapt algorithms, since $\mathcal{M}$ is nonlinear.

- Find is a "good replacement" $+$
- no "global space of directions"; $T_p\mathcal{M} \neq T_q\mathcal{M}$

Task. How to handle the case when f is nonsmooth?

Example application. Given $A, B \in \mathbb{R}^{d\times n}$, $\mathcal{M} = \mathrm{SO}(d)$

- orthogonal Procrustes: $f(p) = \|A - pB\|_{\mathrm{F}}$ has a closed form solution.
- spectral Procrustes: $f(p) = \|A - pB\|_2$ does not

The Riemannian Convex Bundle Method

Hajg Jasa

joint work with
Ronny Bergmann and Roland Herzog

DNA Seminar, NTNU, Trondheim

April 19th, 2024

The Problem

Consider the following minimization problem

$$\text{minimize } f(p), \quad p \in \mathcal{M},$$

The Problem

Consider the following minimization problem

$$\text{minimize } f(p), \quad p \in \mathcal{M},$$

where

- $\mathcal{M}$ is a Riemannian manifold

The Problem

Consider the following minimization problem

$$\text{minimize } f(p), \quad p \in \mathcal{M},$$

where

- $\mathcal{M}$ is a Riemannian manifold
  - complete
  - bounded sectional curvature $\omega \leq \kappa \leq \Omega$ on $\operatorname{dom} f$

NTNU

The Problem

Consider the following minimization problem

$$\text{minimize } f(p), \quad p \in \mathcal{M},$$

where

- $\mathcal{M}$ is a Riemannian manifold
  - complete
  - bounded sectional curvature $\omega \leq \kappa \leq \Omega$ on $\operatorname{dom} f$
- $f\colon \mathcal{M} \to \overline{\mathbb{R}}$ is non-smooth and

The Problem

Consider the following minimization problem

$$\operatorname{minimize} f(p), \quad p \in \mathcal{M},$$

where

- $\mathcal{M}$ is a Riemannian manifold
  - complete
  - bounded sectional curvature $\omega \leq \kappa \leq \Omega$ on $\operatorname{dom} f$
- $f \colon \mathcal{M} \to \overline{\mathbb{R}}$ is non-smooth and
  - $\operatorname{dom} f \neq \emptyset$, $\operatorname{diam}(\operatorname{dom} f) < \infty$ if $\kappa \neq 0$, strongly geodesically convex set and $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$ in $\mathcal{M}$
  - geodesically convex: $f \circ \gamma$ convex
  - lower semi-continuous: $\liminf_{q \to p} f(q) \geq f(p)$

The Problem

Consider the following minimization problem

$$\operatorname{minimize} f(p), \quad p \in \mathcal{M},$$

where

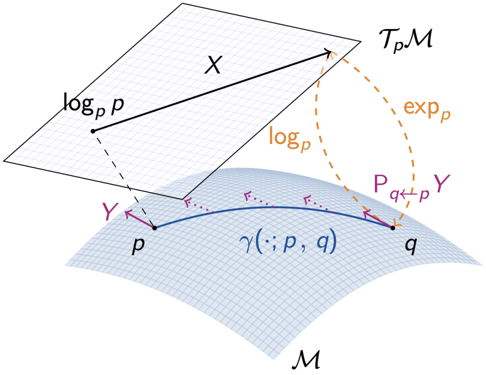
- $\mathcal{M}$ is a Riemannian manifold
  - complete
  - bounded sectional curvature $\omega \leq \kappa \leq \Omega$ on $\operatorname{dom} f$
- $f\colon \mathcal{M} \to \overline{\mathbb{R}}$ is non-smooth and
  - $\operatorname{dom} f \neq \emptyset$, $\operatorname{diam}(\operatorname{dom} f) < \infty$ if $\kappa \neq 0$, strongly geodesically convex set and $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$ in $\mathcal{M}$
  - geodesically convex: $f \circ \gamma$ convex
  - lower semi-continuous: $\liminf_{q \to p} f(q) \geq f(p)$

Goal. Solve this optimization problem with a convex bundle method.

Riemannian Geometry

Notation

- Smooth Riemannian manifold $\mathcal{M}$
- Tangent space $\mathcal{T}_p\mathcal{M}$ at the point $p \in \mathcal{M}$
- Inner product $(\cdot,\cdot)_p : \mathcal{T}_p\mathcal{M} \times \mathcal{T}_p\mathcal{M} \to \mathbb{R}$
- Exponential map $\exp_p X_p = \gamma_{pq}(1) = q$
- Logarithmic map $\log_p q = \exp_p^{-1} q = X_p$
- Parallel transport $\mathrm{P}_{q\leftarrow p} : \mathcal{T}_p\mathcal{M} \to \mathcal{T}_q\mathcal{M}$
- Sectional curvature κ

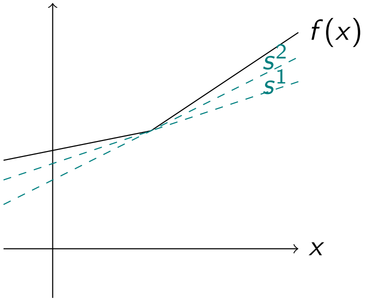


The Convex Subdifferential(s)

For a convex function, the subdifferential is defined as

$$\partial f(x) = \left\{ s \in \mathbb{R}^n \,\middle|\, f(y) \geq f(x) + (s)^T (y - x) \text{ for all } y \in \mathbb{R}^n \right\}$$

and it is a non-empty, closed and convex subset.



The Convex Subdifferential(s)

For a convex function, the subdifferential is defined as

$$\partial f(x) = \left\{ s \in \mathbb{R}^n \,\middle|\, f(y) \geq f(x) + (s)^T (y - x) \text{ for all } y \in \mathbb{R}^n \right\}$$

and it is a non-empty, closed and convex subset.
Let $\varepsilon > 0$. The $\varepsilon-$subdifferential is

$$\partial_\varepsilon f(x) = \left\{ s \in \mathbb{R}^n \,\middle|\, f(y) \geq f(x) + (s)^T (y - x) - \varepsilon \text{ for all } y \in \mathbb{R}^n \right\}$$

and

$$\partial f(x) \subseteq \partial_\varepsilon f(x)$$

The Convex Subdifferential(s)

For a geodesically convex function, the subdifferential is defined as

$$\partial f(p) = \bigl\{ X_p \in \mathcal{T}_p\mathcal{M} \bigm| f(q) \geq f(p) + \bigl(X_p, \log_p q \bigr) \text{ for all } q \in \operatorname{dom} f \bigr\}$$

and it is a non-empty, closed and convex subset.
Let $\varepsilon > 0$. The $\varepsilon-$subdifferential is

$$\partial_\varepsilon f(p) = \bigl\{ X_p \in \mathcal{T}_p\mathcal{M} \bigm| f(q) \geq f(p) + \bigl(X_p, \log_p q \bigr) - \varepsilon \text{ for all } q \in \operatorname{dom} f \bigr\}$$

and

$$\partial f(p) \subseteq \partial_\varepsilon f(p)$$

The Bundle Method

Bundle methods are about descent as well as stability.

The Bundle Method

Bundle methods are about descent as well as stability.

Core characteristics of a general bundle algorithm ([Bonnans et al. 2006]):

The Bundle Method

Bundle methods are about descent as well as stability.

Core characteristics of a general bundle algorithm ([Bonnans et al. 2006]):

- keeps history of the last "best" points

The Bundle Method

Bundle methods are about descent as well as stability.

Core characteristics of a general bundle algorithm ([Bonnans et al. 2006]):

- keeps history of the last "best" points
- solves a stabilization subproblem

The Bundle Method

Bundle methods are about descent as well as stability.

Core characteristics of a general bundle algorithm ([Bonnans et al. 2006]):

- keeps history of the last “best” points
- solves a stabilization subproblem

  dual approach as in [Geiger and Kanzow 2002]

The Bundle Method

Bundle methods are about descent as well as stability.

Core characteristics of a general bundle algorithm ([Bonnans et al. 2006]):

- keeps history of the last “best” points
- solves a stabilization subproblem

  dual approach as in [Geiger and Kanzow 2002]
- generates sequences of *candidate* points and *stability centers*

The Bundle Method

Bundle methods are about descent as well as stability.

Core characteristics of a general bundle algorithm ([Bonnans et al. 2006]):

- keeps history of the last "best" points
- solves a stabilization subproblem

  dual approach as in [Geiger and Kanzow 2002]
- generates sequences of *candidate* points and *stability centers*

Goal. Approximate ∂f with $\partial_\varepsilon f$ on $\operatorname{int}(\operatorname{dom} f)$.

Approximating the $\varepsilon-$subdifferential

Given $x^{(0)}, \ldots, x^{(k)} \in \mathbb{R}^n$, and $s^{(j)} \in \partial f(x^{(j)})$ for $j = 0, \ldots, k$, and

Approximating the $\varepsilon-$subdifferential

Given $x^{(0)}, \ldots, x^{(k)} \in \mathbb{R}^n$, and $s^{(j)} \in \partial f(x^{(j)})$ for $j = 0, \ldots, k$, and

$$e_j^{(k)} := f(x^{(k)}) - f(x^{(j)}) - (s^{(j)})^T(x^{(k)} - x^{(j)}),$$

Approximating the $\varepsilon-$subdifferential

Given $x^{(0)}, \ldots, x^{(k)} \in \mathbb{R}^n$, and $s^{(j)} \in \partial f(x^{(j)})$ for $j = 0, \ldots, k$, and

$$e_j^{(k)} := f(x^{(k)}) - f(x^{(j)}) - (s^{(j)})^T (x^{(k)} - x^{(j)}),$$

then

$$s^{(j)} \in \partial_{e_j^{(k)}} f(x^{(k)}).$$

Approximating the $\varepsilon-$subdifferential

Given $x^{(0)}, \dots, x^{(k)} \in \mathbb{R}^n$, and $s^{(j)} \in \partial f(x^{(j)})$ for $j = 0, \dots, k$, and

$$e_j^{(k)} := f(x^{(k)}) - f(x^{(j)}) - (s^{(j)})^T(x^{(k)} - x^{(j)}),$$

then

$$s^{(j)} \in \partial_{e_j^{(k)}} f(x^{(k)}).$$

Characterize an inner approximation of $\partial_\varepsilon f(p)$ as:

$$G_\varepsilon^{(k)} := \left\{ \sum_{j=0}^{k} \lambda_j s^{(j)} \,\middle|\, \sum_{j=0}^{k} \lambda_j\, e_j^{(k)} \leq \varepsilon,\ \sum_{j=0}^{k} \lambda_j = 1,\ \lambda_j \geq 0\, \forall j = 0, \dots, k \right\}$$

with $G_\varepsilon^{(k)} \subseteq \partial_\varepsilon f(x^{(k)})$.

NTNU

Approximating the $\varepsilon-$subdifferential

Given $x^{(0)}, \ldots, x^{(k)} \in \mathbb{R}^n$, and $s^{(j)} \in \partial f(x^{(j)})$ for $j = 0, \ldots, k$, and

$$e_j^{(k)} := f(x^{(k)}) - f(x^{(j)}) - (s^{(j)})^T(x^{(k)} - x^{(j)}),$$

then

$$s^{(j)} \in \partial_{e_j^{(k)}} f(x^{(k)}).$$

Characterize an inner approximation of $\partial_\varepsilon f(p)$ as:

$$G_\varepsilon^{(k)} := \left\{ \sum_{j=0}^{k} \lambda_j s^{(j)} \,\middle|\, \sum_{j=0}^{k} \lambda_j e_j^{(k)} \leq \varepsilon,\ \sum_{j=0}^{k} \lambda_j = 1,\ \lambda_j \geq 0\, \forall j = 0, \ldots, k \right\}$$

with $G_\varepsilon^{(k)} \subseteq \partial_\varepsilon f(x^{(k)})$.

Main challenge on manifolds: given $p^{(0)}, \ldots, p^{(k)} \in \mathcal{M}$, and $X_{p^{(j)}} \in \partial f(p^{(j)})$, then

$$\mathrm{P}_{p^{(k)} \leftarrow p^{(j)}} X_{p^{(j)}} \in \partial_c f(p^{(k)}) \quad \text{for some} \quad c > 0?$$

Curvature Correction

Using the upper bound Ω on the curvature, define

$$c_j^{(k)} := f(p^{(k)}) - f(p^{(j)}) - \left(X_{p^{(j)}}, \log_{p^{(j)}} p^{(k)}\right) \quad \text{if } \Omega \leq 0,$$
$$c_j^{(k)} := f(p^{(k)}) - f(p^{(j)}) + \|X_{p^{(j)}}\| \|\log_{p^{(j)}} p^{(k)}\| \quad \text{if } \Omega > 0.$$

[Bergmann, Herzog, and HJ 2024].

Curvature Correction

Using the upper bound Ω on the curvature, define

$$c_j^{(k)} := f(p^{(k)}) - f(p^{(j)}) - \Big(X_{p^{(j)}}, \log_{p^{(j)}} p^{(k)} \Big) \quad \text{if } \Omega \leq 0,$$
$$c_j^{(k)} := f(p^{(k)}) - f(p^{(j)}) + \|X_{p^{(j)}}\| \|\log_{p^{(j)}} p^{(k)}\| \quad \text{if } \Omega > 0.$$

[Bergmann, Herzog, and HJ 2024].

We get

$$G_\varepsilon^{(k)} := \left\{ \sum_{j=0}^{k} \lambda_j \mathrm{P}_{p^{(k)} \leftarrow p^{(j)}} X_{p^{(j)}} \;\middle|\; \sum_{j=0}^{k} \lambda_j \, c_j^{(k)} \leq \varepsilon, \ \sum_{j=0}^{k} \lambda_j = 1, \ \lambda_j \geq 0 \text{ for all } j = 0, \dots, k \right\}$$

with $G_\varepsilon^{(k)} \subseteq \partial_\varepsilon f(p^{(k)})$, and $\mathrm{P}_{p^{(k)} \leftarrow p^{(j)}} X_{p^{(j)}} \in \partial_{c_j^{(k)}} f(p^{(k)})$.

The Euclidean Subproblem

Let $k \in \mathbb{N}$ and $j \in \{0, \ldots, k\} = J^{(k)}$ and let λ_j be convex coefficients.

The Euclidean Subproblem

Let $k \in \mathbb{N}$ and $j \in \{0, \ldots, k\} = J^{(k)}$ and let λ_j be convex coefficients.

$$s^{(j)} \in \partial f(x^{(j)}) \implies \sum_{j=0}^{k} \lambda_j s^{(j)} \in \partial_\varepsilon f(x^{(k)})$$

The Euclidean Subproblem

Let $k \in \mathbb{N}$ and $j \in \{0, \ldots, k\} = J^{(k)}$ and let λ_j be convex coefficients.

$$s^{(j)} \in \partial f(x^{(j)}) \implies \sum_{j=0}^{k} \lambda_j s^{(j)} \in \partial_\varepsilon f(x^{(k)}) \quad \text{iff} \quad \varepsilon \geq \sum_{j=0}^{k} \lambda_j e_j^{(k)}$$

The Euclidean Subproblem

Let $k \in \mathbb{N}$ and $j \in \{0, \ldots, k\} = J^{(k)}$ and let λ_j be convex coefficients.

$$s^{(j)} \in \partial f(x^{(j)}) \implies \sum_{j=0}^{k} \lambda_j s^{(j)} \in \partial_\varepsilon f(x^{(k)}) \quad \text{iff} \quad \varepsilon \geq \sum_{j=0}^{k} \lambda_j e_j^{(k)}$$

The *search direction* is

$$d^{(k)} := -\sum_{j=0}^{k} \lambda_j s^{(j)}$$

The Euclidean Subproblem

Let $k \in \mathbb{N}$ and $j \in \{0, \ldots, k\} = J^{(k)}$ and let λ_j be convex coefficients.

$$s^{(j)} \in \partial f(x^{(j)}) \implies \sum_{j=0}^{k} \lambda_j s^{(j)} \in \partial_\varepsilon f(x^{(k)}) \quad \text{iff} \quad \varepsilon \geq \sum_{j=0}^{k} \lambda_j e_j^{(k)}$$

The *search direction* is

$$d^{(k)} := -\sum_{j=0}^{k} \lambda_j s^{(j)}$$

where λ_j are solution to

$$\operatorname*{arg\,min}_{\lambda \in \mathbb{R}^{|J^{(k)}|}} \quad \frac{1}{2} \left\| \sum_{j \in J^{(k)}} \lambda_j s^{(j)} \right\|^2 + \sum_{j \in J^{(k)}} \lambda_j e_j^{(k)}$$

$$\text{s. t.} \quad \sum_{j \in J^{(k)}} \lambda_j = 1, \quad \lambda_j \geq 0 \qquad \text{for all } j \in J^{(k)}$$

The Riemannian Subproblem

Let $k \in \mathbb{N}$ and $j \in \{0, \dots, k\} = J^{(k)}$ and let λ_j be convex coefficients.

$$X_{q^{(j)}} \in \partial f(q^{(j)}) \implies \sum_{j=0}^{k} \lambda_j \mathrm{P}_{p^{(k)} \leftarrow q^{(j)}} X_{q^{(j)}} \in \partial_\varepsilon f(p^{(k)}) \quad \text{iff} \quad \varepsilon \geq \sum_{j=0}^{k} \lambda_j c_j^{(k)}$$

The *search direction* is

$$d^{(k)} := -\sum_{j=0}^{k} \lambda_j \mathrm{P}_{p^{(k)} \leftarrow q^{(j)}} X_{q^{(j)}}$$

where λ_j are solution to

$$\begin{aligned} \operatorname*{arg\,min}_{\lambda \in \mathbb{R}^{|J^{(k)}|}} \quad & \frac{1}{2} \left\| \sum_{j \in J^{(k)}} \lambda_j \mathrm{P}_{p^{(k)} \leftarrow q^{(j)}} X_{q^{(j)}} \right\|^2 + \sum_{j \in J^{(k)}} \lambda_j c_j^{(k)} \\ \text{s. t.} \quad & \sum_{j \in J^{(k)}} \lambda_j = 1, \quad \lambda_j \geq 0 \qquad \text{for all } j \in J^{(k)} \end{aligned}$$

The Euclidean Convex Bundle Method

Data: $x^{(0)} = y^{(0)} \in \mathbb{R}^n$, $g^{(0)} = s^{(0)} \in \partial f(x^{(0)})$, $m \in (0,1)$,
$\varepsilon^{(0)} = e_0^{(0)} = 0$, $k = 0$, $J^{(k)} = \{0\}$

while *not converged* **do**

Compute a solution $\lambda^{(k)} \in \mathbb{R}^{|J^{(k)}|}$ of the subproblem and set

$$g^{(k)} := \sum_{j \in J^{(k)}} \lambda_j^{(k)} s^{(j)}, \quad \varepsilon^{(k)} := \sum_{j \in J^{(k)}} \lambda_j^{(k)} e_j^{(k)},$$

$$d^{(k)} := -g^{(k)}, \quad \xi^{(k)} := -\|g^{(k)}\|^2 - \varepsilon^{(k)}$$

Set $y^{(k+1)} := x^{(k)} + d^{(k)}$.

If $f(y^{(k+1)}) \leq f(x^{(k)}) + m\xi^{(k)}$, then set $x^{(k+1)} := y^{(k+1)}$, else set $x^{(k+1)} := x^{(k)}$.

Compute $s^{(k+1)} \in \partial f(y^{(k+1)})$, update $J^{(k+1)}$, and set $e_j^{(k+1)} := f(x^{(k+1)}) - f(y^{(j)}) - (s^{(j)})^T (x^{(k+1)} - y^{(j)})$ for all $j \in J^{(k+1)}$

end

Result: $x^{(k_*)}$, for some $k_* \in \mathbb{N}$.

The Riemannian Convex Bundle Method

Data: $p^{(0)} = q^{(0)} \in \operatorname{int}(\operatorname{dom} f)$, $g^{(0)} = X_1 \in \partial f(p^{(0)})$, $m \in (0,1)$,
$\varepsilon^{(0)} = c_0^{(0)} = 0$, $k = 0$, $J^{(k)} = \{0\}$, $\beta > 0$, $\Omega \in \mathbb{R}$,

while *not converged* **do**

Compute a solution $\lambda^{(k)} \in \mathbb{R}^{|J^{(k)}|}$ of the subproblem and set

$$g^{(k)} := \sum_{j \in J^{(k)}} \lambda_j^{(k)} \mathrm{P}_{p^{(k)} \leftarrow q^{(j)}} X_{q^{(j)}}, \quad \varepsilon^{(k)} := \sum_{j \in J^{(k)}} \lambda_j^{(k)} c_j^{(k)},$$

$$d^{(k)} := -g^{(k)}, \quad \xi^{(k)} := -\|g^{(k)}\|^2 - \varepsilon^{(k)}$$

Set $t^{(k)} := 1$. While $q^{(k+1)} := \exp_{p^{(k)}}(t^{(k)} d^{(k)}) \notin \operatorname{int}(\operatorname{dom} f)$ or $\operatorname{dist}(q^{(k+1)}, p^{(k)}) < t^{(k)} \|d^{(k)}\|$ backtrack $t^{(k)} = \beta t^{(k)}$.

If $f(q^{(k+1)}) \leq f(p^{(k)}) + m\xi^{(k)}$, then set $p^{(k+1)} := q^{(k+1)}$, else set $p^{(k+1)} := p^{(k)}$.

Compute $X_{q^{(k+1)}} \in \partial f(q^{(k+1)})$, update $J^{(k+1)}$, and set
$c_j^{(k+1)} := f(p^{(k+1)}) - f(q^{(j)}) - \left(X_{q^{(j)}}, \log_{q^{(j)}} p^{(k+1)} \right)$ for all $j \in J^{(k+1)}$ if $\Omega \leq 0$

end

Result: $p^{(k_*)}$, for some $k_* \in \mathbb{N}$.

NTNU

The Riemannian Convex Bundle Method

Data: $p^{(0)} = q^{(0)} \in \operatorname{int}(\operatorname{dom} f)$, $g^{(0)} = X_1 \in \partial f(p^{(0)})$, $m \in (0,1)$,
$\varepsilon^{(0)} = c_0^{(0)} = 0$, $k = 0$, $J^{(k)} = \{0\}$, $\beta > 0$, $\Omega \in \mathbb{R}$,

while *not converged* **do**

Compute a solution $\lambda^{(k)} \in \mathbb{R}^{|J^{(k)}|}$ of the subproblem and set

$$g^{(k)} := \sum_{j \in J^{(k)}} \lambda_j^{(k)} \mathrm{P}_{p^{(k)} \leftarrow q^{(j)}} X_{q^{(j)}}, \quad \varepsilon^{(k)} := \sum_{j \in J^{(k)}} \lambda_j^{(k)} c_j^{(k)},$$

$$d^{(k)} := -g^{(k)}, \quad \xi^{(k)} := -\|g^{(k)}\|^2 - \varepsilon^{(k)}$$

Set $t^{(k)} := 1$. While $q^{(k+1)} := \exp_{p^{(k)}}(t^{(k)} d^{(k)}) \notin \operatorname{int}(\operatorname{dom} f)$ or $\operatorname{dist}(q^{(k+1)}, p^{(k)}) < t^{(k)} \|d^{(k)}\|$ backtrack $t^{(k)} = \beta t^{(k)}$.

If $f(q^{(k+1)}) \leq f(p^{(k)}) + m\xi^{(k)}$, then set $p^{(k+1)} := q^{(k+1)}$, else set $p^{(k+1)} := p^{(k)}$.

Compute $X_{q^{(k+1)}} \in \partial f(q^{(k+1)})$, update $J^{(k+1)}$, and set $c_j^{(k+1)} := f(p^{(k+1)}) - f(q^{(j)}) + \|\log_{q^{(j)}} p^{(k+1)}\| \|X_{q^{(j)}}\|$ for all $j \in J^{(k+1)}$ if $\Omega > 0$

end

Result: $p^{(k_*)}$, for some $k_* \in \mathbb{N}$.

Convergence

- In the Euclidean case [Geiger and Kanzow 2002, Theorem 6.80] holds.

Theorem
Let the solution set $S = \{x^* \in \mathbb{R}^n \mid f(x^*) = \inf_{x \in \mathbb{R}^n} f(x)\}$ *of the minimization problem be nonempty. Then every sequence* $\{x^{(k)}\}$ *generated by the bundle method algorithm converges to a minimizer of* f.

Convergence

- In the Euclidean case [Geiger and Kanzow 2002, Theorem 6.80] holds.

Theorem
Let the solution set $S = \{x^* \in \mathbb{R}^n \mid f(x^*) = \inf_{x\in\mathbb{R}^n} f(x)\}$ *of the minimization problem be nonempty. Then every sequence* $\{x^{(k)}\}$ *generated by the bundle method algorithm converges to a minimizer of* f*.*

- In the Hadamard case, assuming
  1. $t^{(k)} > m$ for all $k \geq k_*$, if a finite number of serious steps k_* occur
  2. no accumulation point of $p^{(k)}$ is allowed to lie on $\partial\,\mathrm{dom}\, f$

  we have an analogous result. [Bergmann, Herzog, and HJ 2024]

Convergence

In [Geiger and Kanzow 2002, Theorem 6.80], they are able to show that

$$\sum_{j=k_*}^{+\infty} \left(\|g^{(j)}\|^2 + \varepsilon^{(j)} \right)^2 < +\infty$$

where $k_* \in \mathbb{N}$ is the index that corresponds to the last serious iterate, namely

$$p^{(k)} = p^{(k_*)} \quad \text{for all } k \geq k_*.$$

This is possible because from the definition of $e_j^{(k)}$ one gets a bound

$$\left(s^{(k)}, g^{(k-1)} \right) < m \left(\|g^{(k-1)}\|^2 + \varepsilon^{(k-1)} \right) - e_k^{(k)},$$

which is still valid in the Hadamard case for $c_k^{(k)}$ with the added assumptions.

Implementation

The algorithm is implemented in Julia using `Manopt.jl` ([Bergmann 2022]) and `Manifolds.jl` ([Axen et al. 2023]). A solver call looks like [1]

```
p* = convex_bundle_method(M, f, ∂f, p0;
     diameter = δ, domain = dom f, k_max = Ω, m = 10^-3)
```

where

- `M` is a Riemannian manifold
- `f` is the objective function
- `∂f` is a subgradient of the objective function
- `p0` is an initial point on the manifold

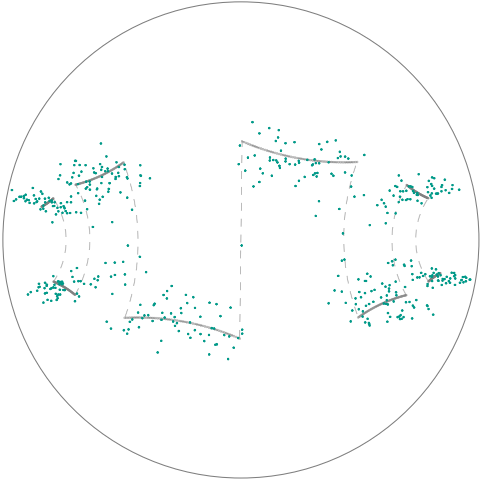
The default stopping criterion for the algorithm is set to

$$-\xi^{(k)} \leq 10^{-8}.$$

[1] https://manoptjl.org/stable/solvers/convex_bundle_method/

Numerical Example: Signal Denoising on $\mathcal{H}^2$

- $\mathcal{H}^2$ hyperbolic space, $n = 496$
- $\mathcal{M} = (\mathcal{H}^2)^n$
- artificial signal $q \in \mathcal{M}$
- noisy signal $\bar{q} \in \mathcal{M}$,
  $\bar{q}^{[i]} = \exp_{q^{[i]}} X_i$, $\sigma = 0.1$



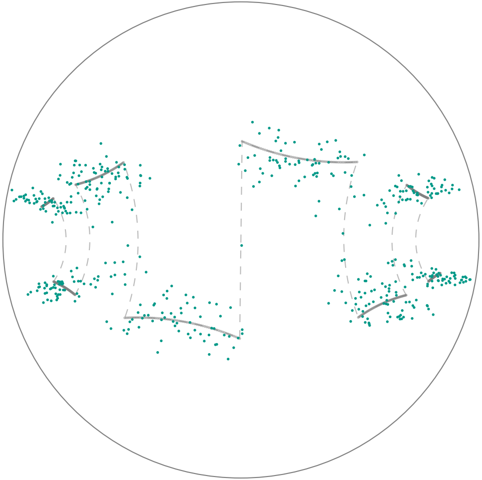
signal q (gray)
noisy signal $\bar{q}$ (teal)

Numerical Example: Signal Denoising on $\mathcal{H}^2$

- $\mathcal{H}^2$ hyperbolic space, $n = 496$
- $\mathcal{M} = (\mathcal{H}^2)^n$
- artificial signal $q \in \mathcal{M}$
- noisy signal $\bar{q} \in \mathcal{M}$,
  $\bar{q}^{[i]} = \exp_{q^{[i]}} X_i$, $\sigma = 0.1$

Solve the total variation problem

- $f_q(p) = \frac{1}{n}\bigl(g(p,q) + \alpha\,\mathrm{TV}(p)\bigr)$
- $g(p,q) = \frac{1}{2}\displaystyle\sum_{i=1}^{n} \operatorname{dist}(p^{[i]}, q^{[i]})^2$
- $\mathrm{TV}(p) = \displaystyle\sum_{i=1}^{n-1} \operatorname{dist}(p^{[i]}, p^{[i+1]})$
- $\alpha = 0.05$,
  $\operatorname{diam}(\operatorname{dom} f) := \texttt{floatmax()} \approx 10^{308}$



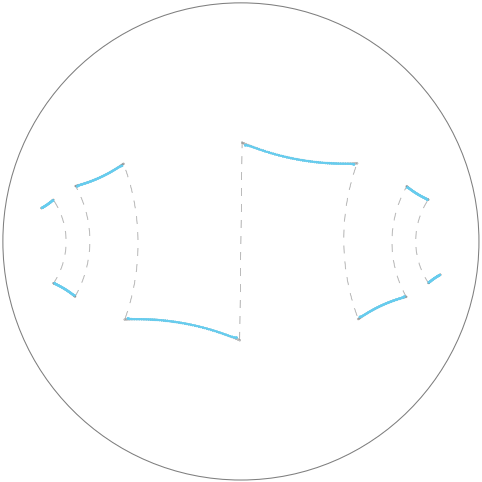
signal q (gray)
noisy signal $\bar{q}$ (teal)

Numerical Example: Signal Denoising on $\mathcal{H}^2$

- $\mathcal{H}^2$ hyperbolic space, $n = 496$
- $\mathcal{M} = (\mathcal{H}^2)^n$
- artificial signal $q \in \mathcal{M}$
- noisy signal $\bar{q} \in \mathcal{M}$,
  $\bar{q}^{[i]} = \exp_{q^{[i]}} X_i$, $\sigma = 0.1$

Solve the total variation problem

- $f_q(p) = \frac{1}{n}\big(g(p,q) + \alpha\,\mathrm{TV}(p)\big)$
- $g(p,q) = \frac{1}{2}\sum_{i=1}^{n} \operatorname{dist}(p^{[i]}, q^{[i]})^2$
- $\mathrm{TV}(p) = \sum_{i=1}^{n-1} \operatorname{dist}(p^{[i]}, p^{[i+1]})$
- $\alpha = 0.05$,
  $\operatorname{diam}(\operatorname{dom} f) := \texttt{floatmax()} \approx 10^{308}$



signal q (gray)
reconstruction q^* (cyan).

Signal Denoising - Algorithms

- Riemannian Convex Bundle Method (RCBM) [Bergmann, Herzog, and HJ 2024]
- Proximal Bundle Algorithm (PBA) [Hoseini Monjezi, Nobakhtian, and Pouryayevali 2021]
- Subgradient Method (SGM) [Ferreira and Oliveira 1998]
- Cyclic Proximal Point Algorithm (CPPA) [Bačák 2014]

Algorithm	Iter.	Time (sec.)	Objective	Error
RCBM	3417	51.393	1.7929×10^{-3}	3.3194×10^{-4}
PBA	15 000	102.387	1.8153×10^{-3}	4.3874×10^{-4}
SGM	15 000	99.604	1.7920×10^{-3}	3.3080×10^{-4}
CPPA	15 000	94.200	1.7928×10^{-3}	3.3230×10^{-4}

Numerical Example: Riemanniann Median on $\mathcal{S}^d$

- $\mathcal{S}^d$ $d-$dimensional sphere
- $\bar{p}$ north pole
- $q^{(1)}, \ldots, q^{(n)} \in \mathcal{S}^d$ are $n = 1000$ Gaussian random data points in $B_{\frac{\pi}{8}}(\bar{p})$
- $\mathcal{D} = \{q^{(1)}, \ldots, q^{(n)} \mid q^{(j)} \in \mathcal{S}^d \text{ for all } j = 1, \ldots, n\}$

Numerical Example: Riemanniann Median on $\mathcal{S}^d$

- $\mathcal{S}^d$ $d-$dimensional sphere
- $\bar{p}$ north pole
- $q^{(1)}, \dots, q^{(n)} \in \mathcal{S}^d$ are $n = 1000$ Gaussian random data points in $B_{\frac{\pi}{8}}(\bar{p})$
- $\mathcal{D} = \{q^{(1)}, \dots, q^{(n)} \mid q^{(j)} \in \mathcal{S}^d \text{ for all } j = 1, \dots, n\}$
- $f(p) = \begin{cases} \frac{1}{n} \sum_{j=1}^{n} \operatorname{dist}(p, q^{(j)}) & \text{if } p \in B_{\frac{\pi}{8}}(\bar{p}), \\ +\infty & \text{otherwise.} \end{cases}$

Numerical Example: Riemanniann Median on $\mathcal{S}^d$

- $\mathcal{S}^d$ $d-$dimensional sphere
- $\bar{p}$ north pole
- $q^{(1)}, \ldots, q^{(n)} \in \mathcal{S}^d$ are $n = 1000$ Gaussian random data points in $B_{\frac{\pi}{8}}(\bar{p})$
- $\mathcal{D} = \{q^{(1)}, \ldots, q^{(n)} \mid q^{(j)} \in \mathcal{S}^d \text{ for all } j = 1, \ldots, n\}$
- $f(p) = \begin{cases} \frac{1}{n}\sum_{j=1}^{n} \operatorname{dist}(p, q^{(j)}) & \text{if } p \in B_{\frac{\pi}{8}}(\bar{p}), \\ +\infty & \text{otherwise.} \end{cases}$

Solve

$$p^* := \operatorname*{arg\,min}_{p \in \mathcal{S}^d} f(p)$$

Riemanniann Median on $\mathcal{S}^d$ - Algorithms

	RCBM			PBA		
Dimension	Iter.	Time (sec.)	Objective	Iter.	Time (sec.)	Objective
2	19	6.50×10^{-3}	0.192 89	20	5.30×10^{-3}	0.192 89
4	28	1.01×10^{-2}	0.198 81	23	5.99×10^{-3}	0.198 81
32	58	2.29×10^{-2}	0.195 76	28	1.13×10^{-2}	0.195 76
1024	48	3.91×10^{-1}	0.197 75	40	3.31×10^{-1}	0.197 75
32 768	43	7.54	0.192 90	21	4.16	0.192 90

	SGM		
Dimension	Iter.	Time (sec.)	Objective
2	5000	1.14	0.192 89
4	3270	8.09×10^{-1}	0.198 81
32	5000	2.18	0.195 76
1024	122	9.75×10^{-1}	0.197 75
32 768	172	5.25×10^{1}	0.192 90

Conclusion and Future Work

In summary:

- introduced the Riemannian Convex Bundle Method for geodesically convex functions on Riemannian manifolds
- discussed convergence and related challenges
- showed two numerical examples

To do:

- further investigate the implications of positive curvature

Selected References

Axen, S. D. et al. (2023). “Manifolds.jl: An Extensible Julia Framework for Data Analysis on Manifolds”. In: *ACM Transactions on Mathematical Software* 49.4. DOI: 10.1145/3618296. arXiv: 2106.08777.

Bačák, M. (2014). “Computing medians and means in Hadamard spaces”. In: *SIAM Journal on Optimization* 24.3, pp. 1542–1566. DOI: 10.1137/140953393.

Bergmann, Ronny (2022). “Manopt.jl: Optimization on Manifolds in Julia”. In: *Journal of Open Source Software* 7.70, p. 3866. DOI: 10.21105/joss.03866.

Bergmann, Ronny, Roland Herzog, and HJ (2024). “The Riemannian Convex Bundle Method”. URL: https://arxiv.org/abs/2402.13670.

Bonnans, J.-F. et al. (2006). *Numerical optimization: theoretical and practical aspects*. Springer-Verlag.

Ferreira, Orizon and Paulo Roberto Oliveira (1998). “Subgradient algorithm on Riemannian manifolds”. In: *Journal of Optimization Theory and Applications* 97.1, pp. 93–104. DOI: 10.1023/A:1022675100677.

Geiger, C. and C. Kanzow (2002). *Theorie und Numerik restringierter Optimierungsaufgaben*. New York: Springer. DOI: 10.1007/978-3-642-56004-0.

Hoseini Monjezi, Najmeh, Soghra Nobakhtian, and Mohamad Reza Pouryayevali (Dec. 2021). “A proximal bundle algorithm for nonsmooth optimization on Riemannian manifolds”. In: *IMA Journal of Numerical Analysis*. DOI: 10.1093/imanum/drab091.