Iteration-Complexity of the Subgradient Method on Riemannian manifolds with lower bounded curvature

Maurício Silva Louzeiro
TU Chemnitz

Based in a joint work with: O. P. Ferreira and L. F. Prudente, both of UFG - Brazil.

Workshop on Optimization on Manifolds
Chemnitz, August 09, 2019
Outline

- Preliminaries

- Iteration-Complexity of the Subgradient Method
  - Subgradient Method with Exogenous Stepsize
  - Subgradient Method with Polyak Stepsize

- Examples
Convexity on Riemannian Manifolds

- $\mathcal{M}$ is a complete Riemannian manifold;
- $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$.
- $f : \mathcal{M} \to \overline{\mathbb{R}}$ is proper, i.e., $\text{dom } f := \{p \in \mathcal{M} : f(p) \neq +\infty\}$ is nonempty.
- $\Gamma_{pq}^f$ is the set of geodesic segments that satisfy $\gamma(0) = p$, $\gamma(1) = q$ and $\gamma(t) \in \text{dom } f$, for all $t \in [0, 1]$.

Definition
A nonempty subset $\Omega \subset \mathcal{M}$ is said to be weakly convex if, for any $p, q \in \Omega$, there is a minimal geodesic segment joining $p$ to $q$ belonging $\Omega$.

Definition
A proper function $f : \mathcal{M} \to \overline{\mathbb{R}}$ is said to be convex on $\mathcal{M}$ if $\text{dom } f$ is weakly convex and

$$f \circ \gamma(t) \leq (1 - t)f(p) + tf(q), \quad \forall \ t \in [0, 1], \ \gamma \in \Gamma_{pq}^f.$$

The subdifferential of a convex function $f : \mathcal{M} \to \mathcal{R}$ at $p \in \text{dom } f$ is defined by

$$\partial f(p) := \left\{ s \in T_p\mathcal{M} : f(q) \geq f(p) + \langle s, \gamma'(0) \rangle, \quad \forall \ q \in \text{dom } f, \ \gamma \in \Gamma_{pq}^f \right\}.$$ 

$\partial f(p)$ is nonempty in all $p \in \text{int dom } f$.

Length comparison result

Let $\mathcal{M}$ with sectional curvature $K \geq \kappa$ such that $\kappa < 0$. Define $\hat{\kappa} := \sqrt{|\kappa|}$.

Lemma

Let $p \in \text{int dom } f$, $0 \neq s \in \partial f(p)$, and let $\gamma : [0, +\infty) \to \mathcal{M}$ be the geodesic defined by $\gamma(t) = \exp_p (-ts/\|s\|)$. Then, for any $t \in [0, \infty)$ and $q \in \text{dom } f$ there holds

$$\cosh(\hat{\kappa} d(\gamma(t), q)) \leq \cosh(\hat{\kappa} d(p, q)) + \hat{\kappa} \cosh(\hat{\kappa} d(p, q)) \sinh(t\hat{\kappa}) \left[ \frac{t}{2} - \frac{\tanh(\hat{\kappa} d(p, q))}{\hat{\kappa} d(p, q)} \frac{f(p) - f(q)}{\|s\|} \right]$$

and, consequently, the following inequality holds

$$d^2(\gamma(t), q) \leq d^2(p, q) + \frac{\sinh (t\hat{\kappa})}{\hat{\kappa}} \left[ \frac{\hat{\kappa} d(p, q)}{\tanh (\hat{\kappa} d(p, q))} t - \frac{2}{\|s\|} (f(p) - f(q)) \right].$$

Lipschitz continuity

Definition
A proper function $f : M \rightarrow \mathbb{R}$ is said to be Lipschitz continuous with constant $\tau \geq 0$ in $\Omega \subset M$ if

$$|f(p) - f(q)| \leq \tau \, d(p, q), \quad p, q \in \Omega.$$
Optimization problem

Consider the optimization problem:

$$\min \{ f(p) : p \in \mathcal{M} \}. \quad (1)$$

- $\mathcal{M}$ is a complete Riemannian manifold with lower bounded curvature $K \geq \kappa$;
- $\Omega^*$ is the solution set of the problem (1);
- $f^* := \inf_{p \in \mathcal{M}} f(p)$ is the optimum value of the problem (1);
- $f : \mathcal{M} \to \mathbb{R}$ is a convex function and lower semicontinuous on $\mathcal{M}$.

The Riemannian subgradient method

Algorithm

**Step 0.** Let \( p_0 \in \text{int} \, \text{dom} \, f \). Set \( k = 0 \).

**Step 1.** If \( s_k = 0 \), then **stop**; otherwise, choose a stepsize \( t_k > 0 \), \( s_k \in \partial f(p_k) \) and compute

\[
p_{k+1} := \exp_{p_k} \left( -t_k \frac{s_k}{\|s_k\|} \right);
\]

**Step 2.** Set \( k \leftarrow k + 1 \) and proceed to **Step 1**.

Assume that \( \{p_k\} \) is well defined and is infinite.
Strategy 1: Exogenous step-size

\[ t_k > 0, \quad \sum_{k=0}^{\infty} t_k = +\infty, \quad \sigma := \sum_{k=0}^{\infty} t_k^2 < +\infty. \]
Strategy 2: Polyak’s step-size

Requirements:

1. $\Omega^* \neq \emptyset$,

2. $d_0 \equiv d(p_0, \Omega^*) = \inf\{d(p_0, q); \ q \in \Omega^*\} > 0$.

$$t_k = \alpha \frac{f(p_k) - f^*}{\|s_k\|}, \quad 0 < \alpha < 2 \frac{\tanh (\hat{\kappa} d_0)}{\hat{\kappa} d_0}.$$ 

Given an estimate $\hat{d} > d_0$ we can choose $0 < \alpha < 2 \tanh (\sqrt{\kappa} \hat{d})/(\sqrt{\kappa} \hat{d})$. 
Exogenous Stepsize: Important inequality

Define

\[ \Omega := \left\{ q \in \mathcal{M} : f(q) \leq \inf_k f(p_k) \right\}. \]

Suppose \( \Omega \neq \emptyset \) and for each \( q \in \Omega \) define

\[ C_{q,\kappa} := \frac{\sinh(\sqrt{\sigma \hat{\kappa}})}{\sqrt{\sigma \hat{\kappa}}} \left[ 1 + \cosh^{-1} \left( \cosh(\hat{\kappa} d(p_0, q)) e^{\frac{1}{2} \sqrt{\sigma \hat{\kappa}} \sinh(\sqrt{\sigma \hat{\kappa}})} \right) \right]. \]

Lemma

If \( \Omega \neq \emptyset \) then, for each \( q \in \Omega \) there holds

\[ d^2(p_{k+1}, q) \leq d^2(p_k, q) + C_{q,\kappa} t_k^2 + 2 \frac{t_k}{\|s_k\|} [f(q) - f(p_k)], \quad s_k \in \partial f(p_k), \quad k = 0, 1, \ldots. \]
Theorem
Assume that $\Omega^* \neq \emptyset$ and $f : \mathcal{M} \to \mathbb{R}$ is Lipschitz continuous with constant $\tau \geq 0$. Then, for all $p* \in \Omega^*$ and every $N \in \mathbb{N}$, the following inequality holds

$$\min \left\{ f(p_k) - f^* : k = 0, 1, \ldots, N \right\} \leq \tau \frac{d^2(p_0, p_*) + C_{p*, \kappa} \sum_{k=0}^{N} t_k^2}{2 \sum_{k=0}^{N} t_k}.$$
Polyak Stepsize: Asymptotic convergence

Theorem

1. \( \liminf_k f(p_k) = f^* \).
2. If \( \Omega^* \neq \emptyset \) then the sequence \( \{p_k\} \) converges to a point \( p_\ast \in \Omega^* \).

Let us define

\[ C_{\kappa,d_0} := \frac{2}{\alpha} - \frac{\hat{\kappa}d_0}{\tanh(\hat{\kappa}d_0)} > 0. \]

**Lemma**

Let \( \bar{q} \in \Omega^* \) such that \( d_0 = d(p_0, \bar{q}) \). Then the following inequality holds

\[ d^2(p_{k+1}, \bar{q}) \leq d^2(p_k, \bar{q}) - C_{\kappa,d_0} \alpha^2 \frac{[f(p_k) - f^*]^2}{\|s_k\|^2}, \quad k = 0, 1, \ldots \]
Polyak Stepsize: Complexity result

**Theorem**

Assume that $f : \mathcal{M} \to \overline{\mathbb{R}}$ is Lipschitz continuous with constant $\tau \geq 0$. Then, for every $N \in \mathbb{N}$, there holds

$$\min \{ f(p_k) - f^* : k = 0, 1, \ldots, N \} \leq \frac{\tau d_0}{\sqrt{C_{\kappa, d_0}}} \frac{1}{\sqrt{(N + 1)}}.$$ 

**Theorem**

The following equality holds $\lim_{k \to \infty} f(p_k) = f^*$. 

Convex feasibility problem

- \( f_i : \mathcal{M} \to \mathbb{R} \) is convex for all \( i = 1, \ldots, m \).
- \( C_i := \{ p \in \mathcal{M} : f_i(p) \leq 0 \} \).

\[
\text{Find } p \in C := \bigcap_{i=1}^{m} C_i, \quad (2)
\]

Consider the problem

\[
\min \{ f(p) : p \in \mathcal{M} \}. \quad (3)
\]

- a solution of (3) with \( f(p) := \max \{ f_1(p), \ldots, f_m(p), 0 \} \) is a solution of (2);
- If \( \text{int} C \neq \emptyset \), exist \( \epsilon > 0 \) such that a solution of (3) with \( f(p) := \max \{ f_1(p), \ldots, f_m(p), -\epsilon \} \) belongs to \( \text{int} C \).

Example family

Consider $\kappa \leq \text{curvature of } \mathcal{M} \leq K$ and set

$$\rho_K := \frac{1}{2} \min \left\{ \text{inj } \mathcal{M}, \frac{\pi}{2\sqrt{K}} \right\},$$

where $\text{inj } \mathcal{M}$ is the injectivity radius of $\mathcal{M}$ and $1/\sqrt{K} := +\infty$ for $K \leq 0$.

Example

Set $q \in \mathcal{M}$, and choose $r > 0$ and $v_1, \ldots, v_m \in T_q \mathcal{M}$ in such a way that

$$a_i := \exp_q \left( r \frac{v_i}{\|v_i\|} \right) \in B(q, \rho), \quad i = 1, \ldots, m,$$

for some $\rho < \rho_K$. Let $\epsilon > 0$ and define $f_i : \mathcal{M} \to \mathbb{R}$ by

$$f_i(p) := d(p, a_i) - r - \epsilon, \quad i = 1, \ldots, m,$$

and consider $f$ given by $f(p) := \max \{ f_1(p), \ldots, f_m(p), -\epsilon \}$. 
Example on SPD matrices

- $\mathbb{P}^n$ - cone of symmetric matrices;
- $\mathbb{P}^{n+}$ - cone of symmetric positive definite matrices.

Let $\mathcal{M} := (\mathbb{P}^{n+}, \langle \cdot, \cdot \rangle)$ be the Riemannian manifold with

$$\langle U, V \rangle := \text{tr}(VX^{-1}UX^{-1}),$$

for all $X \in \mathcal{M}, U, V \in T_X \mathcal{M} \approx \mathbb{P}^n$.

- Exogenous step-size.
Example on unit sphere

Let $S := \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the $(n - 1)$-dimensional unit sphere. Consider $S$ with the Euclidean metric $\langle \cdot, \cdot \rangle$. The tangent plane at $x \in S$ is given by

$$T_x S := \{v \in \mathbb{R}^n : \langle v, x \rangle = 0\}.$$

- Polyak’s step-size.
References


Thank you for your attention!