



HAL
open science

On Approximating the Riemannian 1-Center

Marc Arnaudon, Frank Nielsen

► **To cite this version:**

| Marc Arnaudon, Frank Nielsen. On Approximating the Riemannian 1-Center. 2011. hal-00560187v1

HAL Id: hal-00560187

<https://hal.science/hal-00560187v1>

Preprint submitted on 27 Jan 2011 (v1), last revised 30 Jan 2012 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On Approximating the Riemannian 1-Center

Marc Arnaudon^a, Frank Nielsen^{b,c,*}

^a*Laboratoire de Mathématiques et Applications
CNRS: UMR 6086, Université de Poitiers
Téléport 2 - BP 30179*

F-86962 Futuroscope Chasseneuil Cedex, France

^b*Ecole Polytechnique*

Computer Science Department (LIX)

Palaiseau, France.

^c*Sony Computer Science Laboratories, Inc. (FRL),*

3-14-13 Higashi Gotanda 3F, Shinagawa-Ku,

Tokyo 141-0022, Japan.

Abstract

In this paper, we generalize the simple Euclidean 1-center approximation algorithm of Bădoiu and Clarkson (2003) to Riemannian geometries and study accordingly the convergence rate. We then show how to instantiate this generic algorithm to two particular cases: (1) hyperbolic geometry, and (2) Riemannian manifold of symmetric positive definite matrices.

Keywords: 1-center; minimax ; circumcenter ; Riemannian geometry ; core-set ; approximation

1. Introduction and prior work

Finding the unique smallest enclosing ball (SEB) of a finite Euclidean point set $P = \{p_1, \dots, p_n\}$ is a fundamental problem that has been thoroughly investigated by the computational geometry community Welzl (1991); Nielsen and Nock (2009). This problem is also known in the literature as the minimum enclosing ball (MEB), the 1-center problem, or the minimax

*Corresponding author

Email addresses: Marc.Arnaudon@math.univ-poitiers.fr (Marc Arnaudon),

Frank.Nielsen@acm.org (Frank Nielsen)

<http://www.informationgeometry.org/RiemannMinimax/>

optimization problem in operations research. In practice, since computing exactly the SEB is intractable in high dimensions, efficient approximation algorithms have been proposed. An algorithmic breakthrough was recently achieved by Bădoiu and Clarkson (2008) that proved the existence of a *core-set* $C \subseteq P$ of optimal size $\lceil \frac{1}{\epsilon} \rceil$ so that $r(C) \leq (1+\epsilon)r(P)$ (for arbitrary $\epsilon > 0$), where $r(S)$ denote the radius of the SEB of S . Let $c(S)$ denote the ball center, ie. the circumcenter. Since the size of the core-set depends *only* on the approximation precision ϵ and is independent of the dimension, core-sets have become popular in high-dimensional applications such as supervised classification in machine learning (eg., core vector machines Tsang et al. (2007)). In Bădoiu and Clarkson (2003), a fast and simple approximation algorithm is designed as follows:

BC-ALG:

Starts with $c_1 \in P$ and iteratively update the current center using the rule $c_{i+1} = c_i + \frac{f_i - c_i}{i+1}$, where f_i denotes the farthest point of P to c_i .

It can be proved that a $(1+\epsilon)$ -approximation of the SEB is obtained after $\lceil \frac{1}{\epsilon^2} \rceil$ iterations, thereby showing the existence of a core-set $C = \{f_1, f_2, \dots\}$ of a size at most $\lceil \frac{1}{\epsilon^2} \rceil$: $r(C) \leq (1+\epsilon)r(P)$. This simple algorithm run in time $O(\frac{dn}{\epsilon^2})$, and has been generalized to Bregman divergences Nock and Nielsen (2005) that includes the (squared) Euclidean distance, and are the canonical distances of flat spaces including the particular case of Euclidean geometry. (Note that if we start from the optimal center $c_1 = c(S)$, the first iteration yields a center c_2 away from $c(S)$.)

Many data-sets arising in medical imaging Pennec (2008) or in computer vision Turaga and Chellappa (2010) cannot be considered as emanating from vectorial spaces but rather as lying on curved manifolds. For example, the space of rotations or the space of invertible matrices are not flat as the arithmetic average of two elements does not necessarily lie inside the space.

In this work, we extend the Euclidean BC-ALG algorithm to Riemannian geometries. In the remainder, we assume the Reader familiar to basic notions of Riemannian geometries (refer to Berger (2003) for an introductory textbook) in order not to burden the paper with technical Riemannian definitions. However in appendix 6 we recall some specific notions which play a key role in the paper as geodesics, sectional curvature, injectivity radius,

Alexandrov and Toponogov theorems, and cosine laws for triangles. Furthermore, we consider probability measures instead of finite points sets² so as to study the most general setting.

Let M be a complete Riemannian manifold and ν a probability measure on M . Denote by $\rho(x, y)$ the Riemannian metric distance from x to y on M . Assume the measure support $\text{supp}(\nu)$ is included in a geodesic ball $B(o, R)$. Let

$$R_{\alpha,p} = \begin{cases} \frac{1}{2} \min \left\{ \text{inj}(M), \frac{\pi}{2\alpha} \right\} & \text{if } 1 \leq p < 2 \\ \frac{1}{2} \min \left\{ \text{inj}(M), \frac{\pi}{\alpha} \right\} & \text{if } 2 \leq p \leq \infty \end{cases} \quad (1)$$

where $\text{inj}(M)$ is the injectivity radius (see appendix 6).

For $p \in [1, \infty]$, under the assumption that

$$R < R_{\alpha,p} \quad (2)$$

Afsari proved Afsari (2011) that there exists a *unique* point c_p which minimizes the following cost function

$$\begin{aligned} H_p &: M \rightarrow [0, \infty] \\ x &\mapsto \|\rho(x, \cdot)\|_{L^p(\nu)} \end{aligned} \quad (3)$$

with $c_p \in B(o, R)$ (in fact, lying inside the closure of the convex hull of the masses). Recall that if $p \in [1, \infty)$ then

$$\|\rho(x, \cdot)\|_{L^p(\nu)} = \left(\int_M \rho(x, y)^p \nu(dy) \right)^{1/p}$$

and

$$\|\rho(x, \cdot)\|_{L^\infty(\nu)} = \min \{a > 0, \nu(\{y \in M, \rho(x, y) > a\}) = 0\}.$$

For discrete uniform measure viewed as a ‘‘point cloud’’ and $p \in [1, \infty)$ the map H_p translates as $(\frac{1}{n} \sum_{i=1}^n \|p_i - x\|_p^p)^{1/p}$, with $\|\cdot\|_p$ denoting the L_p norm, and $H_\infty(x)$ is the distance from x to its farthest point in the cloud. The point c_p that realizes the minimum represents a notion of centrality of

²We view finite point sets as discrete uniform probability measures.

the measure (eg., median for $p = 1$, mean for $p = 2$, and circumcenter for $p \rightarrow \infty$). This center is a *global* minimizer (not only in $B(o, R)$), and this explains why a bound for the sectional curvature is required on the whole manifold M (in fact $B(o, 2R)$ is sufficient, see Afsari (2011)).

Deterministic algorithms (subgradient algorithms) for finding c_p have been considered in Yang (2009) for the median case ($p = 1$). Stochastic algorithms have been investigated in Arnaudon et al. (2010) for the case $p \in [1, \infty)$, and a central limit theorem (CLT) is derived (in fact a kind of invariance principle).

In this work, we consider the case $p = \infty$, with c_∞ denoting the circumcenter. In this case there is no canonical deterministic algorithm which generalizes the gradient descent algorithms considered for $p \in [1, \infty)$. Following Eq. 3, $H_\infty(x)$ denotes the farthest distance of x to the measure (L_∞ -norm).

To give an example of a Riemannian manifold, consider for example the space of symmetric positive definite matrices with associated Riemannian distance (see Section 4)

$$\rho(P, Q) = \|\log(P^{-1}Q)\|_F = \sqrt{\sum_i \log^2 \lambda_i} \quad (4)$$

where λ_i are the eigenvalues of matrix $P^{-1}Q$. This is a non-compact Riemannian symmetric space of nonpositive curvature (Cartan-Hadamard manifold, see Lang (1999), chapter 12). In this context *any* measure ν with *bounded support* satisfies assumption Eq. 2 (since we can take $\alpha > 0$ as small as one likes), and consequently the minimizer c_∞ of H_∞ exists and is unique. We call it the 1-center or minimax center of ν .

We generalized the BC-ALG by noticing that the iterative update is a rewriting barycenter rule of the current circumcenter with the current farthest point. Thus the new position of the circumcenter falls along the straight line joining these two points in Euclidean geometry. In Riemannian geometry, the shortest path linking two points is called a geodesic (eg., great arc for spherical geometry). Instead of walking on the straight line, we rather then walk on the geodesic to the furthest point as follows:

GEO-ALG:

Starts with $c_1 \in P$ and iteratively update the current circumcenter as follows: $c_{i+1} = \text{Geodesic}(c_i, f_i, \frac{1}{i+1})$, where f_i denotes the farthest point of P to c_i , and $\text{Geodesic}(p, q, t)$ denotes the intermediate point m on the geodesic passing through p and q such that $\rho(p, m) = t \times \rho(p, q)$.

Note that GEO-ALG generalized BC-ALG by taking the Euclidean distance $\rho(p, q) = \|p - q\|$.

The paper is organized as follows: Section 2 proves technically a crucial lemma. It is followed by the description and convergence rate analysis of our generic Riemannian algorithm in Section 3. Section 4 instantiates the algorithm for the particular cases of the hyperbolic manifold and the manifold of symmetric positive definite matrices. Section 5 concludes the paper and hints at further perspectives. To make the paper self-contained, Section 6 recalls the fundamental notions of Riemannian geometry used throughout the paper.

2. A key lemma

In this section, we assume³ that $\text{supp}(\nu) \subset B(o, R)$ and

$$R < R_{\alpha, \infty} = \frac{1}{2} \min \left\{ \text{inj}(M), \frac{\pi}{\alpha} \right\}$$

The following lemma is essential for proving the convergence of the algorithm determining the minimax of ν .

Lemma 1. *There exists $\tau > 0$ such that for all $x \in B(o, R)$,*

$$H_{\infty}(x) - H_{\infty}(c_{\infty}) \geq \tau \rho^2(x, c_{\infty}). \quad (5)$$

Proof:

The point c_{∞} is the center of the smallest ball which contains $\text{supp}(\nu)$ and the radius of this ball is exactly $r^* := H_{\infty}(c_{\infty})$ (see Afsari (2009)). Denoting by $S(c_{\infty}, r^*)$ the boundary of this ball and by $S_{c_{\infty}}M$ the set of unitary vectors in $T_{c_{\infty}}M$, for all $v \in S_{c_{\infty}}M$ there exists $y \in S(c_{\infty}, r^*) \cap \text{supp}(\nu)$ such that

$$\langle \dot{\gamma}_0(c_{\infty}, y), v \rangle \leq 0 \quad (6)$$

³Any bounded measure on a Cartan-Hadamard manifold satisfies this assumption.

where $t \mapsto \gamma_t(c_\infty, y)$ is the geodesic from c_∞ to y in time one and $\dot{\gamma}_t(c_\infty, y)$ denotes derivative with respect to t . Indeed, if this was not true it would contradict the minimality of $S(c_\infty, r^*)$ (Afsari (2009)).

Now letting $t \mapsto \gamma_t(v) = \exp_x(tv)$ the geodesic satisfying $\dot{\gamma}_0(v) = v$, we prove Eq. 5 for $x = \gamma_t(v)$. We have

$$H_\infty(\gamma_t(v)) - H_\infty(c_\infty) \geq \rho(\gamma_t(v), y) - \rho(c_\infty, y) = \rho(\gamma_t(v), y) - r^* \quad (7)$$

by definition of H_∞ .

Then we consider a 2-dimensional sphere $S_{\alpha^2}^2$ with *constant* curvature α^2 , distance function $\tilde{\rho}$, and in $S_{\alpha^2}^2$ a *comparison triangle* $\tilde{\gamma}_t(\tilde{v})\tilde{y}\tilde{c}_\infty$ such that $\tilde{\rho}(\tilde{y}, \tilde{c}_\infty) = r^*$, \tilde{v} is a unitary vector in $T_{\tilde{c}_\infty}M_\alpha$ satisfying

$$\langle \dot{\tilde{\gamma}}_0(\tilde{c}_\infty, \tilde{y}), \tilde{v} \rangle = \langle \dot{\gamma}_0(c_\infty, y), v \rangle \quad (8)$$

Let us prove that

$$\tilde{\rho}(\tilde{\gamma}_t(\tilde{v}), \tilde{y}) - r^* = \tilde{\rho}(\tilde{\gamma}_t(\tilde{v}), \tilde{y}) - \rho(\tilde{c}_\infty, \tilde{y}) \geq \tau_\alpha \tilde{\rho}^2(\tilde{\gamma}_t(\tilde{v}), \tilde{c}_\infty) \quad (9)$$

for some $\tau_\alpha > 0$ provided condition Eq. 6 is realized: using the first law of cosines (Theorem 4 in appendix), we get

$$0 \geq \cos(\dot{\tilde{\gamma}}_0(\tilde{c}_\infty, \tilde{y}), \tilde{v}) = \frac{\cos(\alpha \tilde{\rho}(\tilde{\gamma}_t(v), \tilde{y})) - \cos(\alpha r^*) \cos(\alpha t)}{\sin(\alpha r^*) \sin(\alpha t)} \quad (10)$$

which yields

$$\cos(\alpha \tilde{\rho}(\tilde{\gamma}_t(\tilde{v}), \tilde{y})) \leq \cos(\alpha r^*) \cos(\alpha t)$$

and this in turn implies

$$\sin(\alpha(\tilde{\rho}(\tilde{\gamma}_t(\tilde{v}), \tilde{y}) - \alpha r^*)) \geq \cotan(\alpha r^*) (\cos(\alpha(\tilde{\rho}(\tilde{\gamma}_t(\tilde{v}), \tilde{y}) - \alpha r^*)) - \cos(\alpha t))$$

so

$$\liminf_{t \searrow 0} \frac{\tilde{\rho}(\tilde{\gamma}_t(\tilde{v}), \tilde{y}) - r^*}{t^2} \geq \frac{\alpha}{2} \cotan(\alpha r^*) \geq \frac{\alpha}{2} \cotan(\alpha R_{\alpha, \infty})$$

uniformly in \tilde{v} . Consequently Eq. 9 is true in a neighborhood of \tilde{c}_∞ , and by a compactness argument we prove that it is true in any compact included in $\bar{B}(\tilde{c}_\infty, R_{\alpha, \infty})$.

To finish the proof we are left to use Alexandrov comparison theorem (Theorem 2 in appendix) with triangles $\gamma_t(v)yc_\infty$ and $\tilde{\gamma}_t(\tilde{v})\tilde{y}\tilde{c}_\infty$ to check that the right hand side of Eq. 7 in M is larger than the left hand side of Eq. 9. This proves Eq. 5 in $B(c_\infty, R) \cap B(o, R)$, and for proving it in $B(o, R)$ we just have to notice that H_∞ is continuous and positive on the compact set $\bar{B}(o, R) \setminus B(c_\infty, R)$, hence it has a positive lower bound. \square

3. Riemannian approximation algorithm

For $x \in B(o, R)$, denote by $t \mapsto \gamma_t(v(x, \nu))$ a unit speed geodesic from $\gamma_0(v(x, \nu)) = x$ to one point $y = \gamma_{H_\infty(x)}(v(x, \nu))$ in $\text{supp}(\nu)$ which realizes the maximum of the distance from x to $\text{supp}(\nu)$. So $v = \frac{1}{H_\infty(x)} \exp_x^{-1}(y)$. A measurable choice is always possible. But note that if ν has finite support, it is natural to make a random choice.

We consider the following stochastic algorithm.

RIE-ALG:

Fix some $\delta > 0$.

Step 1 Choose a starting point $x_0 \in \text{supp}(\nu)$ and let $k = 0$

Step 2 Choose a step size $t_k \in (0, \delta]$ and let $x_{k+1} = \gamma_{t_k}(v(x_k, \nu))$, then do again step 2 with $k = k + 1$.

This algorithm generalizes the Euclidean scheme Bădoiu and Clarkson (2003) (and algorithm GEO-ALG for probability measures). Let $a \wedge b$ denote the minimum operator $a \wedge b = \min(a, b)$.

$$\text{Let } R_0 = \frac{R_{\alpha, \infty} - R}{2} \wedge \frac{R}{2}.$$

Theorem 1. *Assume $\alpha, \beta > 0$ are such that $-\beta^2$ is a lower bound and α^2 an upper bound of the sectional curvatures in M .*

If the step sizes $(t_k)_{k \geq 1}$ verify

$$\delta \leq \frac{R_0}{2} \wedge \frac{2}{\beta} \operatorname{arctanh}(\tanh(\beta R_0/2) \cos(\alpha R) \tan(\alpha R_0/4)), \quad (11)$$

$$\lim_{k \rightarrow \infty} t_k = 0, \quad \sum_{k=0}^{\infty} t_k = +\infty \quad \text{and} \quad \sum_{k=0}^{\infty} t_k^2 < \infty. \quad (12)$$

then the sequence $(x_k)_{k \geq 1}$ generated by the algorithm satisfies

$$\lim_{k \rightarrow \infty} \rho(x_k, c_\infty) = 0. \quad (13)$$

Proof:

First we prove that for all $r \in [R_0, R]$, if $x_k \in B(c_\infty, r)$ then $x_{k+1} \in B(c_\infty, r)$: if $\rho(x_k, c_\infty) \leq R_0/2$ it is clear since $\delta \leq R_0/2$. If $\rho(x_k, c_\infty) \geq$

$R_0/2$ we prove that $\rho(x_{k+1}, c_\infty) \leq \rho(x_k, c_\infty)$. Let $y_{k+1} = \gamma_{H_\infty(x_k)}(v(x_k, \nu))$: $y_{k+1} \in \text{supp}(\nu)$ is such that $H_\infty(x_k) = \rho(x_k, y_{k+1})$; consider the triangle $c_\infty x_k y_{k+1}$. Let $a = \rho(x_k, y_{k+1})$, $b = \rho(y_{k+1}, c_\infty)$ and $r = \rho(c_\infty, x_k)$, \hat{x}_k the angle corresponding to the point x_k . By Alexandrov comparison theorem (in fact Corollary 1 in appendix) \hat{x}_k is smaller than the same in constant curvature α^2 . This together with the law of cosines in spherical geometry (Theorem 4 in appendix) yields

$$\cos \hat{x}_k \geq \frac{\cos \alpha b - \cos \alpha r \cos \alpha a}{\sin \alpha r \sin \alpha a}.$$

Now $r \geq R_0/2$, $b \leq r^*$ and $a \geq r^*$ so

$$\cos \hat{x}_k \geq \frac{\cos \alpha r^*(1 - \cos(\alpha R_0/2))}{\sin(\alpha R_0/2)} = \cos \alpha r^* \tan(\alpha R_0/4) \geq \cos \alpha R \tan(\alpha R_0/4). \quad (14)$$

Consider now the triangle $c_\infty x_k x_{k+1}$ and let $f = \rho(c_\infty, x_{k+1})$. Recall $\rho(x_k, x_{k+1}) = t_{k+1}$. Now by Toponogov theorem (Theorem 3 in appendix) f is smaller than the same in constant curvature $-\beta^2$. This together with first law of cosines in hyperbolic geometry (Theorem 4 in appendix) yields

$$\cosh \beta f \leq \cosh \beta r \cosh \beta t_{k+1} - \cos \hat{x}_k \sinh \beta r \sinh \beta t_{k+1} \quad (15)$$

which implies by Eq. 14

$$\cosh \beta f \leq \cosh(\beta r) \cosh \beta t_{k+1} - \cos \alpha R \tan(\alpha R_0/4) \sinh(\beta r) \sinh \beta t_{k+1} \quad (16)$$

and we easily check that the condition on δ implies that the right hand side is smaller than $\cosh \beta r$. This proves that $\rho(c_\infty, x_{k+1}) \leq \rho(c_\infty, x_k)$.

Then we prove that there exists $\eta > 0$ such that if $x_k \in B(c_\infty, R) \setminus B(c_\infty, R_0)$ then

$$\frac{\cosh(\beta \rho(c_\infty, x_{k+1}))}{\cosh(\beta \rho(c_\infty, x_k))} \leq 1 - \eta t_{k+1}. \quad (17)$$

From Eq. 16, we obtain

$$\begin{aligned} \frac{\cosh \beta f}{\cosh \beta r} &\leq \cosh \beta t_{k+1} - \cos \alpha R \tan(\alpha R_0/4) \tanh(\beta r) \sinh \beta t_{k+1} \\ &\leq \cosh \beta t_{k+1} - \cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0) \sinh \beta t_{k+1} \\ &\leq 1 - 2(\cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0) \cosh(\beta t_{k+1}/2)) \end{aligned}$$

$$\begin{aligned}
& -\sinh(\beta t_{k+1}/2) \sinh(\beta t_{k+1}/2) \\
\leq & 1 - (\cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0) \cosh(\beta t_{k+1}/2) - \sinh(\beta t_{k+1}/2)) \beta t_{k+1} \\
\leq & 1 - (\cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0) \\
& - \cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0/2)) \cosh(\beta t_{k+1}/2) \beta t_{k+1}
\end{aligned}$$

where we used Eq. 11 in the last inequality. So

$$\begin{aligned}
\frac{\cosh \beta \rho(c_\infty, x_{k+1})}{\cosh \beta \rho(c_\infty, x_k)} & \leq 1 - (\cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0) \\
& - \cos \alpha R \tan(\alpha R_0/4) \tanh(\beta R_0/2)) \beta t_{k+1}
\end{aligned} \tag{18}$$

and this gives Eq. 17, using the fact that $x \mapsto \cosh \beta x$ has derivative bounded below by $\beta \sinh \beta R_0$ on $[R_0, R]$.

At this stage, since $\sum_{k=1}^{\infty} t_k = \infty$, we can conclude that there exists k_0 such that $\cosh(\beta \rho(c_\infty, x_{k_0})) \leq \cosh(\beta R_0)$ so $x_{k_0} \in B(c_\infty, R_0)$, and for all $k \geq k_0$, $x_k \in B(c_\infty, R_0)$.

Now we use the fact that on $B(c_\infty, R_0)$, H_∞ is convex and satisfies Eq. 5. By boundedness of the Hessian of square distance to c_∞ (see Yang (2009) for details), we have for $k \geq k_0$

$$\begin{aligned}
\rho^2(c_\infty, x_{k+1}) & \leq \\
\rho^2(c_\infty, x_k) - 2t_{k+1} \langle \exp_{x_k}^{-1} c_\infty, \dot{\gamma}_0(v(x_k, \nu)) \rangle & + C \left(\frac{R_{\alpha, \infty} + R}{2}, \beta \right) t_{k+1}^2
\end{aligned} \tag{19}$$

with

$$C(r, \beta) = 2r\beta \cotanh(2\beta r). \tag{20}$$

Now letting $y_{k+1} = \gamma_{H_\infty(x_k)}(v(x_k, \nu))$ we have $H_\infty \geq \rho(\cdot, y_{k+1})$ since $y_{k+1} \in \text{supp}(\nu)$. we remark that $\rho^2(\cdot, y_{k+1})$ is convex on $B(c_\infty, R_0)$ by the fact that for all $z \in B(c_\infty, R_0)$ and $y \in \text{supp}(\nu)$, $\rho(z, y) < R_{\alpha, \infty}$. Moreover we have $H_\infty(x_k) = \rho(x_k, y_{k+1})$. As a consequence, we get

$$\begin{aligned}
H_\infty(c_\infty) - H_\infty(x_k) & \geq \rho^2(c_\infty, y_{k+1}) - \rho^2(x_k, y_{k+1}) \\
& \geq -2 \langle \exp_{x_k}^{-1} c_\infty, \dot{\gamma}_0(v(x_k, \nu)) \rangle
\end{aligned}$$

and this implies by Proposition 1

$$-2 \langle \exp_{x_k}^{-1} c_\infty, \dot{\gamma}_0(v(x_k, \nu)) \rangle \leq -\tau \rho^2(c_\infty, x_k). \quad (21)$$

Plugging into Eq. 19 yields

$$\rho^2(c_\infty, x_{k+1}) \leq (1 - \tau t_{k+1}) \rho^2(c_\infty, x_k) + C \left(\frac{R_{\alpha, \infty} + R}{2}, \beta \right) t_{k+1}^2. \quad (22)$$

We recall from here the standard argument to prove that $\rho^2(c_\infty, x_k)$ converges to 0. Let

$$a = \limsup_{k \rightarrow \infty} \rho^2(c_\infty, x_k).$$

Iterating Eq. 22 yields for $\ell \geq 1$

$$\rho^2(c_\infty, x_{k+\ell}) \leq \prod_{j=1}^{\ell} (1 - \tau t_{k+j}) \rho^2(c_\infty, x_k) + C \sum_{j=1}^{\ell} t_{k+j}^2$$

with $C = C \left(\frac{R_{\alpha, \infty} + R}{2}, \beta \right)$. Letting $\ell \rightarrow \infty$ and using the fact that $\sum_{j=1}^{\infty} t_{k+j} = \infty$, which implies

$$\prod_{j=1}^{\infty} (1 - \tau t_{k+j}) = 0,$$

we get

$$a \leq C \sum_{j=1}^{\infty} t_{k+j}^2.$$

Finally using $\sum_{j=1}^{\infty} t_j^2 < \infty$ we obtain that $\lim_{k \rightarrow \infty} \sum_{j=1}^{\infty} t_{k+j}^2 = 0$, so $a = 0$. □

For the speed of convergence, taking $t_k = \frac{r}{k+1}$, we proceed as in Proposition 4.10 in Yang (2009). We use the following lemma, borrowed from Nedic and Bertsekas (2000):

Lemma 2. *Let $(u_k)_{k \geq 1}$ be a sequence of nonnegative real numbers such that*

$$u_{k+1} \leq \left(1 - \frac{\lambda}{k+1} \right) u_k + \frac{\xi}{(k+1)^2}$$

where λ and ξ are positive constants. Then

$$u_{k+1} \leq \begin{cases} \frac{1}{(k+1)^\lambda} \left(u_0 + \frac{2^\lambda \xi (2-\lambda)}{1-\lambda} \right) & \text{if } 0 < \lambda < 1; \\ \frac{\xi(1+\ln(k+1))}{k+1} & \text{if } \lambda = 1; \\ \frac{1}{(\lambda-1)(k+2)} \left(\xi + \frac{(\lambda-1)u_0 - \xi}{(k+2)^{\lambda-1}} \right) & \text{if } \lambda > 1. \end{cases}$$

Choosing $t_k = \frac{r}{k+1}$, letting k_0 such that for all $k \geq k_0$, $x_k \in B(c_\infty, R_0)$,

$$\rho^2(x_{k_0+k}, c_\infty) \leq \begin{cases} \frac{1}{(k+1)^\lambda} \left(R_0^2 + \frac{2^\lambda \xi (2-\lambda)}{1-\lambda} \right) & \text{if } 0 < \lambda < 1; \\ \frac{\xi(1+\ln(k+1))}{k+1} & \text{if } \lambda = 1; \\ \frac{1}{(\lambda-1)(k+2)} \left(\xi + \frac{(\lambda-1)R_0^2 - \xi}{(k+2)^{\lambda-1}} \right) & \text{if } \lambda > 1. \end{cases}$$

where $\lambda = \tau r$ (with τ given in proposition 1) and $\xi = r^2 C \left(\frac{R_{\alpha, \infty} + R}{2}, \beta \right)$.

Proof:

This is a direct consequence of lemma 2 and inequality Eq. 22, valid for $k \geq k_0$. \square

Remark 1. It is possible to obtain an estimate of η in Eq. 17, and from this one can get an estimate of k_0 .

Remark 2. We can choose R_0 as small as we want, and as $R_0 \rightarrow 0$, one can let $\tau \rightarrow \frac{\alpha}{2} \cotan(\alpha R_{\alpha, \infty})$. Again explicit estimates are possible.

4. Two case studies

In order to implement algorithm GEO-ALG (a specialization of RIE-ALG for point clouds with step sizes $t_i = \frac{1}{i+1}$), we need to describe the geodesics of the underlying manifold, and find an intermediate point $m = \text{Geodesic}(p, q, t)$ on the geodesic passing through p and q such that $\rho(p, m) = t \rho(p, q)$.

4.1. Hyperbolic manifold

A hyperbolic manifold is a complete Riemannian d -dimensional manifold of constant sectional curvature -1 that is isometric to the real hyperbolic space. There exists several models of hyperbolic geometry. Here, we consider the planar Klein model where geodesics are straight lines Nielsen and Nock

(2010). Although there exists no known closed-form formula for the hyperbolic centroid ($p = 2$), Welzl's minimax algorithm generalizes to the Klein disk Nielsen and Nock (2010) to compute exactly the hyperbolic circumcenter. The Klein Riemannian distance is

$$\rho(p, q) = \operatorname{arccosh} \frac{1 - p^t q}{\sqrt{(1 - p^t p)(1 - q^t q)}} \quad (23)$$

where $\operatorname{arccosh}(x) = \log(x + \sqrt{x^2 - 1})$, and the geodesic passing through p and q is the straight line segment

$$\gamma_t(p, q) = (1 - t)p + tq \quad (24)$$

Finding m such that $\rho(p, m) = t \rho(p, q)$ cannot be solved in closed-form solution (except for $t = \frac{1}{2}$, see Nielsen and Nock (2010)), so that we rather proceed by a bisection search algorithm up to machine precision.

4.2. Manifold of symmetric positive definite matrices

A $d \times d$ matrix M with real entries is said symmetric positive definite (SPD) iff. it is symmetric ($M = M^T$), and that for all $x \neq 0$, $x^T M x > 0$. The set of $d \times d$ SPD matrices form a smooth manifold of dimension $\frac{d(d+1)}{2}$. We refer to Lang (1999) (Chapter 12) for a description of the geometry of SPD matrices. See also Ji (2007) for optimization on matrix manifolds. The geodesic linking (matrix) point P to point Q is given by

$$\gamma_t(P, Q) = P^{\frac{1}{2}} \left(P^{-\frac{1}{2}} Q P^{-\frac{1}{2}} \right)^t P^{\frac{1}{2}}, \quad (25)$$

where the matrix function $h(M)$ is computed from the singular value decomposition $M = U D V^T$ (with U and V unitary matrices and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_d)$ a diagonal matrix of eigenvalues) as $h(M) = U \operatorname{diag}(h(\lambda_1), \dots, h(\lambda_d)) V^T$. For example, the square root function of a matrix is computed as $M^{\frac{1}{2}} = U \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_d}) V^T$.

In this case, finding t such that

$$\|\log(P^{-1}Q)^t\|_F^2 = r \|\log P^{-1}Q\|_F^2, \quad (26)$$

where $\|\cdot\|_F$ denotes the Fröbenius norm yields to $t = r$. Indeed, consider $\lambda_1, \dots, \lambda_d$ the eigenvalues of $P^{-1}Q$, then Eq. 26 amounts to find

$$\sum_{i=1}^d \log^2 \lambda_i^t = t^2 \sum_{i=1}^d \log^2 \lambda_i = r^2 \sum_{i=1}^d \log^2 \lambda_i. \quad (27)$$

That is $t = r$.

5. Concluding remarks and discussion

We described a generalization of the 1-center algorithm of Bădoiu and Clarkson (2003) to arbitrary Riemannian geometry, and proved the convergence under mild assumptions. This proves the existence of Riemannian core-sets for optimization. This 1-center building block can be used for k -center clustering. Furthermore, the algorithm can be straightforwardly extended to sets of geodesic balls.

A source code implementation in JAVATM is available at

<http://www.informationgeometry.org/RiemannMinimax/>

Acknowledgements

The second author thanks Prasenjit Saha for several discussions related to this topic, and gratefully acknowledge financial support from French funding agency ANR (contract GAIA 07-BLAN-0328-01) and Sony Computer Science Laboratories, Inc.

6. Appendix: Some notions of Riemannian geometry

In this section, we recall some basic notions of Riemannian geometry used throughout the paper. For a complete presentation, we refer to Cheeger and Ebin (1975).

We let M be a Riemannian manifold and $\langle \cdot, \cdot \rangle$ the Riemannian metric, which is a definite positive bilinear form on each tangent space $T_x M$, and depends smoothly on x . The associated norm in $T_x M$ will be denoted by $\|\cdot\|$: $\|u\| = \langle u, u \rangle^{1/2}$. We denote by $\rho(x, y)$ the distance between two points on the manifold M :

$$\rho(x, y) = \inf \left\{ \int_0^1 \|\dot{\varphi}(t)\| dt, \varphi \in C^1([0, 1], M), \varphi(0) = x, \varphi(1) = y \right\}.$$

A *geodesic* in M is a smooth path which locally minimizes the distance between two points. In general such a curve does not minimize it globally. However it is true in all the sets we are considering in this paper. Given a vector $v \in TM$ with base point x , there is a unique geodesic started at x with speed v at time 0. It is denoted by $t \mapsto \exp_x(tv)$ or compactly by $t \mapsto \gamma_t(v)$. It depends smoothly on v but it has in general finite lifetime. A geodesic defined on a time interval $[a, b]$ is said to be *minimal* if it minimizes the distance from the image of a to the image of b . If the manifold is complete, taking $x, y \in M$, there exists a minimal geodesic from x to y in time 1. In all the scenarii we are considering in this paper, the minimal geodesic is unique and depends smoothly on x and y , and we denote it by $\gamma.(x, y) : [0, 1] \rightarrow M$, $t \mapsto \gamma_t(x, y)$ with the conditions $\gamma_0(x, y) = x$ and $\gamma_1(x, y) = y$. A subset U of M is said to be *convex* if for any $x, y \in U$, there exists a unique minimal geodesic $\gamma.(x, y)$ in M from x to y , this geodesic fully lies in U and depends smoothly on x, y, t .

The *injectivity radius* of M , denoted by $\text{inj}(M)$, is the largest $r > 0$ such that for all $x \in M$, the map \exp_x restricted to the open ball in T_xM centered at 0 with radius r is an embedding.

Given $x \in M$, u, v two non collinear vectors in T_xM , the *sectional curvature* $\text{Sect}(u, v) = K$ is a number which gives information on how the geodesics issued from x behave near x . More precisely the image by \exp_x of the circle centered at 0 of radius $r > 0$ in $\text{Span}(u, v)$ has length

$$2\pi S_K(r) + o(r^3) \quad \text{as } r \rightarrow 0$$

with

$$S_K(r) = \begin{cases} \frac{\sin(\sqrt{K}r)}{\sqrt{K}} & \text{if } K > 0, \\ r & \text{if } K = 0, \\ \frac{\sinh(\sqrt{-K}r)}{\sqrt{-K}} & \text{if } K < 0. \end{cases}$$

For instance, if $K > 0$, $\exp_x(\text{Span}(u, v))$ is near x approximatively a 2-dimensional sphere with radius $\frac{1}{\sqrt{K}}$. In fact, if M is simply connected and all the sectional curvatures are equal to the same $K > 0$, then M is a d -dimensional sphere with radius $\frac{1}{\sqrt{K}}$, where d is the dimension of M . If M is simply connected and all the sectional curvatures are equal to the same $K < 0$, we say that M is a d -dimensional hyperbolic space with curvature K .

An upper bound (resp. lower bound) of sectional curvatures is a number a such that for all non collinear u, v in the same tangent space, $\text{Sect}(u, v) \leq a$

(resp. $\text{Sect}(u, v) \geq a$). In the paper, we used a positive upper bound α^2 and a negative lower bound $-\beta^2$, $\alpha, \beta > 0$.

The existence of the upper bound α^2 for sectional curvatures makes possible to compare geodesic triangles, by *Alexandrov* theorem (see Chavel (2003)).

Theorem 2. *Let $x_1, x_2, x_3 \in M$ satisfy $x_1 \neq x_2$, $x_1 \neq x_3$ and*

$$\rho(x_1, x_2) + \rho(x_2, x_3) + \rho(x_3, x_1) < 2 \min \left\{ \text{inj}M, \frac{\pi}{\alpha} \right\}$$

where $\alpha > 0$ is such that α^2 is an upper bound of sectional curvatures. Let the minimizing geodesic from x_1 to x_2 and the minimizing geodesic from x_1 to x_3 make an angle θ at x_1 . Denoting by $S_{\alpha^2}^2$ the 2-dimensional sphere of constant curvature α^2 (hence of radius $1/\alpha$) and $\tilde{\rho}$ the distance in $S_{\alpha^2}^2$, we consider points $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3 \in S_{\alpha^2}^2$ such that $\rho(x_1, x_2) = \tilde{\rho}(\tilde{x}_1, \tilde{x}_2)$, $\rho(x_1, x_3) = \tilde{\rho}(\tilde{x}_1, \tilde{x}_3)$. Assume that the minimizing geodesic from \tilde{x}_1 to \tilde{x}_2 and the minimizing geodesic from \tilde{x}_1 to \tilde{x}_3 also make an angle θ at \tilde{x}_1 .

Then we have $\rho(x_2, x_3) \geq \tilde{\rho}(\tilde{x}_2, \tilde{x}_3)$.

Instead of prescribing the angle in the comparison triangle in the sphere, it is possible to prescribe the third distance:

Corollary 1. *The assumption are the same as in Theorem 2 except that we assume that $\rho(x_2, x_3) = \tilde{\rho}(\tilde{x}_2, \tilde{x}_3)$ (all the distances are equal), but the minimizing geodesic from \tilde{x}_1 to \tilde{x}_2 and the minimizing geodesic from \tilde{x}_1 to \tilde{x}_3 now make an angle $\tilde{\theta}$ at \tilde{x}_1 .*

Then we have $\tilde{\theta} \geq \theta$.

There also exists a comparison result in the other direction, called *Topogonov's* theorem.

Theorem 3. *Assume $\beta > 0$ is such that $-\beta^2$ is a lower bound for sectional curvatures in M . Let $x_1, x_2, x_3 \in M$ satisfy $x_1 \neq x_2$, $x_1 \neq x_3$. Let the minimizing geodesic from x_1 to x_2 and the minimizing geodesic from x_1 to x_3 make an angle θ at x_1 . Denoting by $H_{-\beta^2}^2$ the hyperbolic 2-dimensional space of constant curvature $-\beta^2$ and $\tilde{\rho}$ the distance in $H_{-\beta^2}^2$, we consider points $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3 \in H_{-\beta^2}^2$ such that $\rho(x_1, x_2) = \tilde{\rho}(\tilde{x}_1, \tilde{x}_2)$, $\rho(x_1, x_3) = \tilde{\rho}(\tilde{x}_1, \tilde{x}_3)$. Assume that the minimizing geodesic from \tilde{x}_1 to \tilde{x}_2 and the minimizing geodesic from \tilde{x}_1 to \tilde{x}_3 also make an angle θ at \tilde{x}_1 .*

Then we have $\rho(x_2, x_3) \leq \tilde{\rho}(\tilde{x}_2, \tilde{x}_3)$.

Triangles in the sphere $S_{\alpha^2}^2$ and in the hyperbolic space $H_{-\beta^2}^2$ have explicit relations between distance and angles as we will see below. This combined with Theorems 2 and 3 and Corollary 1 allow to find related bounds in M , which are intensively used in our proofs.

In this paper, we only use the *first law of cosines* in $S_{\alpha^2}^2$ and in $H_{-\beta^2}^2$ (see e.g. Ratcliffe (1994) Theorem 2.5.3 and Theorem 3.5.3).

Theorem 4. *If $\theta_1, \theta_2, \theta_3$ are the angles of a triangle in $S_{\alpha^2}^2$ and x_1, x_2, x_3 are the lengths of the opposite sides, then*

$$\cos \theta_3 = \frac{\cos(\alpha x_3) - \cos(\alpha x_1) \cos(\alpha x_2)}{\sin(\alpha x_1) \sin(\alpha x_2)}.$$

If $\theta_1, \theta_2, \theta_3$ are the angles of a triangle in $H_{-\beta^2}^2$ and x_1, x_2, x_3 are the lengths of the opposite sides, then

$$\cos \theta_3 = \frac{\cosh(\beta x_1) \cosh(\beta x_2) - \cosh(\beta x_3)}{\sinh(\beta x_1) \sinh(\beta x_2)}.$$

References

- Afsari, B., 2009. Means and averaging on Riemannian manifolds. Ph.D. thesis, University of Maryland.
- Afsari, B., February 2011. Riemannian L^p center of mass : existence, uniqueness, and convexity. Proceedings of the American Mathematical Society 139, 655–674.
- Arnaudon, M., Dombry, C., Phan, A., Yang, L., 2010. Stochastic algorithms for computing means of probability measures. arXiv, Hal-00540623.
- Berger, M., 2003. A panoramic view of Riemannian geometry. Springer Verlag, Berlin.
- Bădoiu, M., Clarkson, K. L., 2003. Smaller core-sets for balls. In: SODA '03: Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, pp. 801–802.
- Bădoiu, M., Clarkson, K. L., May 2008. Optimal core-sets for balls. Computational Geometry: Theory and Applications 40, 14–22.

- Chavel, I., 2006. Riemannian geometry: A modern introduction. Cambridge University Press, 2nd edition, 2006.
- Cheeger, J., Ebin, D.G., 1975. Comparison Theorems in Riemannian Geometry. North-Holland mathematical library, Vol. 9.
- Ji, H., 2007. Optimization approaches on smooth manifolds. PhD thesis, Australian National University.
- Lang, S., 1999. Fundamentals of differential geometry. Vol. 191 of Graduate Texts in Mathematics. Springer-Verlag, New York.
- Nedic, A., Bertsekas, D., 2000. Convergence rate of incremental subgradient algorithms. In: Stochastic Optimization: Algorithms and Applications. Kluwer, pp. 263–304.
- Nielsen, F., Nock, R., 2009. Approximating smallest enclosing balls with applications to machine learning. *Int. J. Comput. Geometry Appl.* 19 (5), 389–414.
- Nielsen, F., Nock, R., 2010. Hyperbolic Voronoi diagrams made easy. In: International Conference on Computational Science and its Applications (ICCSA). Vol. 0. IEEE Computer Society, Los Alamitos, CA, USA, pp. 74–80.
- Nock, R., Nielsen, F., 2005. Fitting the smallest enclosing bregman ball. In: European Conference on Machine Learning (ECML). pp. 649–656.
- Turaga, P., Veeraraghavan, A., Srivastava, A., Chellappa, R., 2010. Statistical computations on Grassmann and Stiefel manifolds for image and video based recognition. *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)*.
- Penneç, X., 2008. Statistical computing on manifolds: From riemannian geometry to computational anatomy. In: Emerging Trends in Visual Computing (ETVC). pp. 347–386.
- Ratcliffe, J., 1994. Foundations of hyperbolic manifolds. Graduate texts in Mathematics, Springer-Verlag, 1994.
- Tsang, I. W., Kocsor, A., Kwok, J. T., 2007. Simpler core vector machines with enclosing balls. In: ICML '07: Proceedings of the 24th international conference on Machine learning. ACM, New York, NY, USA, pp. 911–918.

Welzl, E., 1991. Smallest enclosing disks (balls and ellipsoids). In: Maurer, H. (Ed.), *New Results and New Trends in Computer Science*. LNCS. Springer.

Yang, L., 2009. Riemannian Median and Its Estimation. arXiv, Hal-00432906.