Consensus on Lie groups for the Riemannian Center of Mass

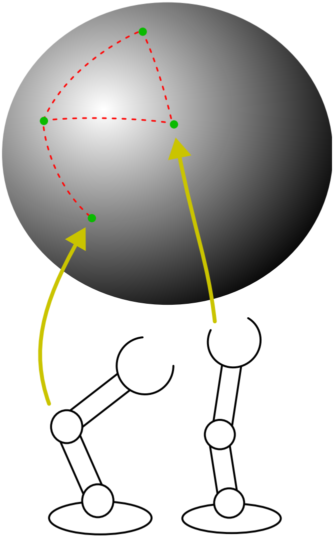
Presenter: Spencer Kraisler
Coauthors: Shahriar Talebi and Mehran Mesbahi

RAIN Lab
William E. Boeing Department of Aeronautics & Astronautics
University of Washington

December 2023

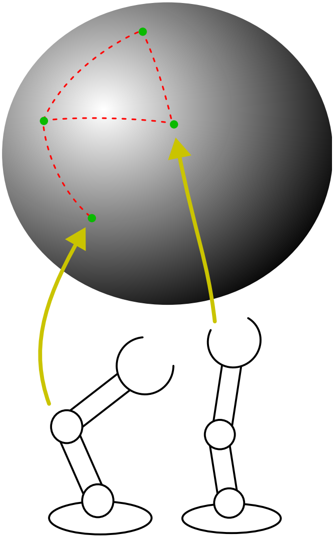
Motivation

- **Example**: 3D localization via network of cameras [Tron, 2012]
- **Example**: Coordinated motion of robot arms [Sarlette, 2010]
    - **Non-Euclidean** state space (Dome camera, covariance matrix, $SO(3)$, robot arm)



Motivation

- **Example**: 3D localization via network of cameras [Tron, 2012]
- **Example**: Coordinated motion of robot arms [Sarlette, 2010]
  - **Non-Euclidean** state space (Dome camera, covariance matrix, $SO(3)$, robot arm)
- Consensus is the **foundation** of distributed computation
  - Synchronize states of network of processors $\iff$ steer agents to **single point**
- Consensus point needs geometric+statistical **significance**



Background

- **Smooth manifold**:
  - topological space $\mathcal{M} \subset \mathbb{R}^n$
  - locally Euclidean
  - every point has a tangential space $T_x\mathcal{M}$
- **Riemannian metric**:
  - Geodesic distance: $d_g(.,.)$
  - $\exp_x : T_x\mathcal{M} \to \mathcal{M}$ and $\log_x(.) := (\exp_x(.))^{-1}$
- intrinsic vs. extrinsic quantities

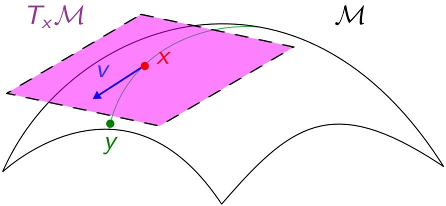


Figure: Start at x, travel along v for $\|v\|$ distance, arrive at $y := \exp_x(v)$

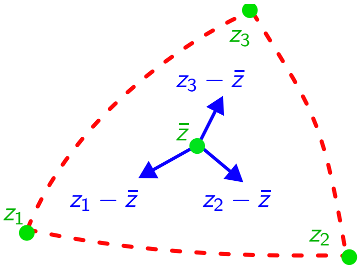
Example

	Euclidean	$SO(n)$	**Riemannian**
Point	$x \in \mathbb{R}^n$	$R \in SO(n)$	$x \in \mathcal{M}$
Tangent vectors	$v \in \mathbb{R}^n$	$W \in T_R SO(n) = R \cdot \text{Skew}(n)$	$v \in T_x\mathcal{M}$
Inner product	$v^T w$	$\text{trace}(V^T W)$	$\langle v, w \rangle_x$
Geodesic	$\gamma_{x,v}(t) = x + tv$	$\gamma_{R,V}(t) = R \exp(tV)$	$\gamma_{x,v}(t) = \exp_x(tv)$

What is an average on a Riemannian manifold?

Euclidean:

$$\bar{x} := \arg\min_{x \in \mathbb{R}^n} \sum_{i=1}^{N} \|x - z_i\|^2$$

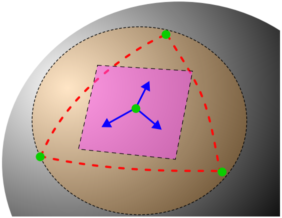


Background II

What is an average on a Riemannian manifold?

Euclidean:

$$\bar{x} := \arg\min_{x \in \mathbb{R}^n} \sum_{i=1}^{N} \|x - z_i\|^2$$



Riemannian center of mass (RCM):

$$\textbf{RCM}(z_1, \ldots, z_N) := \arg\min_{x \in \mathcal{M}} \sum_{i=1}^{N} d_g(x, z_i)^2$$

Applications: medical imaging, averaging correlation matrices, averaging attitudes for attitude estimation

Background II

What is an average on a Riemannian manifold?
Euclidean:

$$\bar{x} := \arg\min_{x \in \mathbb{R}^n} \sum_{i=1}^{N} \|x - z_i\|^2$$

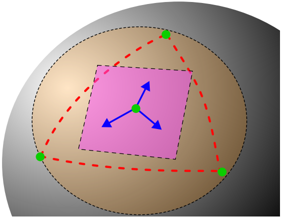
Riemannian center of mass (RCM):

$$\mathbf{RCM}(z_1, \ldots, z_N) := \arg\min_{x \in \mathcal{M}} \sum_{i=1}^{N} d_g(x, z_i)^2$$

Applications: medical imaging, averaging correlation matrices, averaging attitudes for attitude estimation

Instrinsic better than extrinsic:

$$\bar{z}_{ext} := \mathrm{Proj}_{\mathcal{M}}\left(\frac{1}{N}\sum_{i=1}^{N} z_i\right)$$



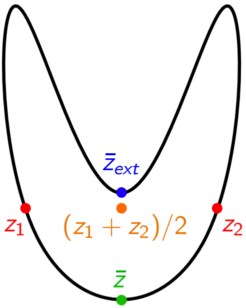


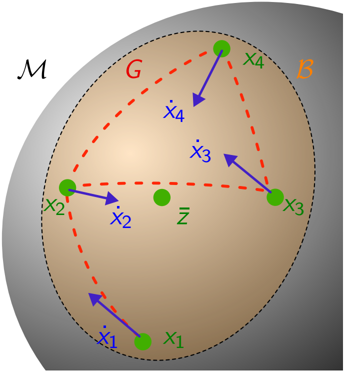
Figure: Flaw of extrinsic mean

Problem Formulation

Consider:

- Complete Riemannian manifold $\mathcal{M}$
- $z_1, \ldots, z_N \in \mathcal{B} \subset \mathcal{M}$ (only available locally)
- Agents/processors $x_1, \ldots, x_N \in \mathcal{M}$ under connected communication network $G = ([N], E)$
- Distributed dynamics:

$$\dot{x}_i(t) = \mathbf{F}_i(x_j(t) : j \in \mathcal{N}_i \cup \{i\})$$

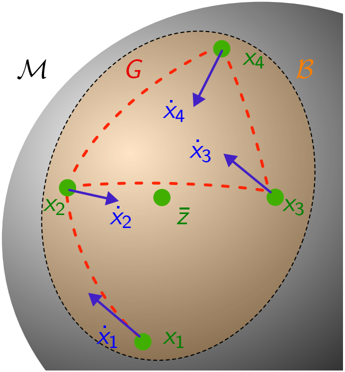


Problem Formulation

Consider:

- Complete Riemannian manifold $\mathcal{M}$
- $z_1, \ldots, z_N \in \mathcal{B} \subset \mathcal{M}$ (only available locally)
- Agents/processors $x_1, \ldots, x_N \in \mathcal{M}$ under connected communication network $G = ([N], E)$
- Distributed dynamics:

$$\dot{x}_i(t) = \mathbf{F}_i(x_j(t) : j \in \mathcal{N}_i \cup \{i\})$$



Problem

Design intrinsic distributed dynamics $\mathbf{F}_i$ *such that*

$$\lim_{t \to \infty} x_i(t) = \mathbf{RCM}(z_1, \ldots, z_N)$$

Naive solution

Euclidean:
$\dot{x}_i = \sum_{j \in i}(x_j - x_i)$

Theorem

Under the above dynamics, consensus is guaranteed and each $x_i(t) \rightarrow x^* = \frac{1}{N}\sum_{i=1}^{N} x_i(0)$.

[1]Tron [2013]
[2]Sarlette [2008]

Naive solution

Euclidean:

$\dot{x}_i = \sum_{j \in i}(x_j - x_i)$

Theorem

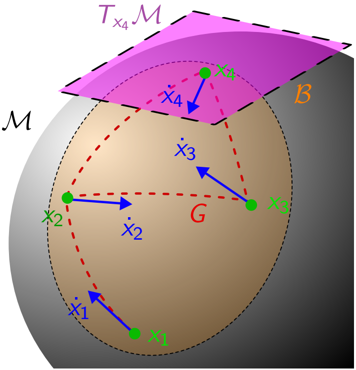
Under the above dynamics, consensus is guaranteed and each $x_i(t) \rightarrow x^* = \frac{1}{N}\sum_{i=1}^{N} x_i(0)$.

Riemannian:

(1) $\dot{x}_i = \sum_{j \in \mathcal{N}_i} \log_{x_i}(x_j)$

Theorem

Under (1), consensus is guaranteed, provided $\{x_i(0)\}$ *are initialized within a geodesic ball of radius* $r < r^*$.



[1] Tron [2013]
[2] Sarlette [2008]

Naive solution

Euclidean:
$\dot{x}_i = \sum_{j \in i}(x_j - x_i)$

Theorem

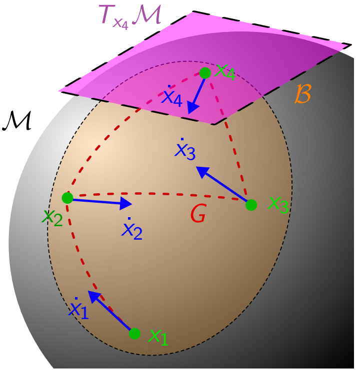
Under the above dynamics, consensus is guaranteed and each $x_i(t) \to x^* = \frac{1}{N}\sum_{i=1}^{N} x_i(0)$.

Riemannian:

1. $\dot{x}_i = \sum_{j \in \mathcal{N}_i} \log_{x_i}(x_j)$
2. **or** $\dot{x}_i = \mathcal{P}_{T_{x_i}\mathcal{M}}\left[\sum_{j \in \mathcal{N}_i}(x_j - x_i)\right]$

Theorem

Under (1), consensus is guaranteed, provided $\{x_i(0)\}$ *are initialized within a geodesic ball of radius* $r < r^*$.



[1] Tron [2013]
[2] Sarlette [2008]

Distributed Gradient Descent (DGD) with Gradient Tracking (GT)

Problem: Let G be a connected graph, and each node is assigned a local cost $f_i : \mathbb{R}^n \to \mathbb{R}$. Minimize in distributed fashion

$$f(x) := \sum_{i=1}^{N} f_i(x).$$

Distributed Gradient Descent (DGD) with Gradient Tracking (GT)

Problem: Let G be a connected graph, and each node is assigned a local cost $f_i : \mathbb{R}^n \to \mathbb{R}$. Minimize in distributed fashion

$$f(x) := \sum_{i=1}^{N} f_i(x).$$

DGD:

$$\dot{x}_i = -\nabla f_i(x_i) + \sum_{j \in \mathcal{N}_i} (x_j - x_i)$$

Distributed Gradient Descent (DGD) with Gradient Tracking (GT)

Problem: Let G be a connected graph, and each node is assigned a local cost $f_i : \mathbb{R}^n \to \mathbb{R}$. Minimize in distributed fashion

$$f(x) := \sum_{i=1}^N f_i(x).$$

DGD with GT:

$$\begin{cases} \dot{x}_i = -v_i + \sum_{j \in \mathcal{N}_i}(x_j - x_i) \\ \dot{w}_i = \sum_{j \in \mathcal{N}_i}(v_i - v_j) \\ v_i = -w_i + \nabla f_i(x_i) \\ v_i(0) := 0 \end{cases}$$

Idea: $v_i \approx \frac{1}{N} \sum_{i=1}^N \nabla f_i(x_i)$

Thm [Carnevale, 2023]: Under strong convexity and smoothness assumption, each x_i approaches x^* with **exponential rate**, where x^* minimizies f.

Distributed Gradient Descent (DGD) with Gradient Tracking (GT)

Problem: Let G be a connected graph, and each node is assigned a local cost $f_i : \mathbb{R}^n \to \mathbb{R}$. Minimize in distributed fashion

$$f(x) := \sum_{i=1}^{N} f_i(x).$$

DGD with GT:

$$\begin{cases} \dot{x}_i = -v_i + \sum_{j \in \mathcal{N}_i} (x_j - x_i) \\ \dot{w}_i = \sum_{j \in \mathcal{N}_i} (v_i - v_j) \\ v_i = -w_i + \nabla f_i(x_i) \\ v_i(0) := 0 \end{cases}$$

Idea: $v_i \approx \frac{1}{N} \sum_{i=1}^{N} \nabla f_i(x_i)$

Thm [Carnevale, 2023]: Under strong convexity and smoothness assumption, each x_i approaches x^* with **exponential rate**, where x^* minimizies f.

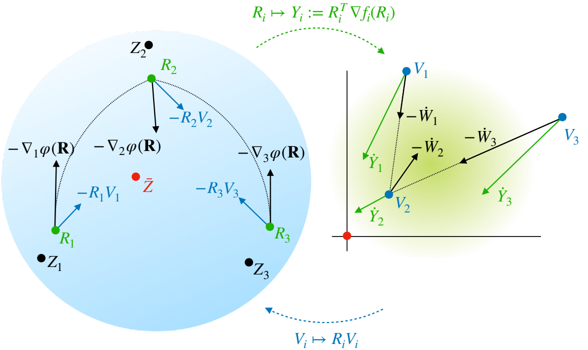
Note: $RCM(z_1, ..., z_N)$ solves $\min_{x \in \mathcal{M}} \sum_{i=1}^{N} d_g(x, z_i)^2$

Generalization of DGD with GT to $SO(3)$

Algorithm ($SO(3)$)

Let $Z_1, \ldots, Z_N \in \mathcal{B} \subset SO(3)$*. Define* $f_i(R) := d_g^2(R, Z_i) = \|\log(R^T Z_i)\|_F^2$ *and initialize* $R_i(0) \in \mathcal{B}$ *and* $V_i(0) := 0$*. Our algorithm follows:*

$$\begin{cases} \dot{R}_i = R_i \cdot \left[-V_i + \sum_{j\in\mathcal{N}_i} \log\left(R_i^T R_j\right)\right] \\ \dot{W}_i = \sum_{j\in\mathcal{N}_i} (V_i - V_j) \\ V_i = -W_i + R_i^T \nabla f_i(R_i) \end{cases}$$

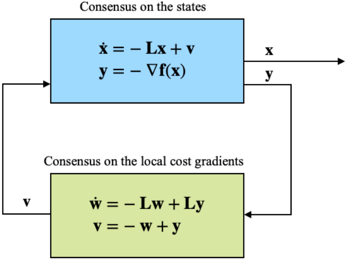


Results

Algorithm (Riemannian manifold)

Let $z_1, \ldots, z_N \in \mathcal{B} \subset \mathcal{M}$. *Define* $f_i(x) := d_g(x, z_i)^2$ *and initialize* $x_i(0) \in \mathcal{B}$ *and* $v_i(0) := 0 \in T_{z_i}\mathcal{M}$. *Our algorithm follows:*

$$\begin{cases} \dot{x}_i = -\mathcal{T}_{z_i}^{x_i} v_i + \sum_{j \in \mathcal{N}_i} \log_{x_i}(x_j) \\ \dot{w}_i = \sum_{j \in \mathcal{N}_i} (v_i - v_j) \\ v_i = -w_i + \mathcal{T}_{x_i}^{z_i} \nabla f_i(x_i) \end{cases}$$



Theorem

Suppose $\mathcal{M}$ *is a Lie group equipped with a bi-invariant metric. Let* $\mathcal{B} \subset \mathcal{M}$ *be a geodesically convex ball with* $z_1, \ldots, z_N \in \mathcal{B}$. *Then the only stationary point of the proposed dynamics in* $\mathcal{B}$ *is* $\mathbf{x}^* = (\bar{z}, \ldots, \bar{z}) \in \mathcal{M}^N$.

Corollary

If $\mathcal{M} = \mathbb{R}^n$, *this stationary point is globally asymptotically stable.*

Proof overview

Let $\mathcal{M}$ be a Lie group with bi-invariant metric. Set $\mathfrak{m} := T_I\mathcal{M}$. Let $(x_1(.), \ldots, x_N(.)) = \mathbf{x}(.) \in \mathcal{B}^N \subset \mathcal{M}^N$ and $(w_1(.), \ldots, w_N(.)) = \mathbf{w}(.) \in \mathfrak{m}^N$.

Proof overview

Let $\mathcal{M}$ be a Lie group with bi-invariant metric. Set $\mathfrak{m} := T_I\mathcal{M}$. Let $(x_1(.), \ldots, x_N(.)) = \mathbf{x}(.) \in \mathcal{B}^N \subset \mathcal{M}^N$ and $(w_1(.), \ldots, w_N(.)) = \mathbf{w}(.) \in \mathfrak{m}^N$.

Prop. 1. Define **geodesic consensus error** as

$$\varphi(\mathbf{x}) := \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} d_g(x_i, x_j)^2$$

Then

$$\nabla\varphi(\mathbf{x}) = 0 \wedge \mathbf{x} \in \mathcal{B}^N \iff \varphi(\mathbf{x}) = 0 \iff x_1 = \cdots = x_N$$

Proof overview

Let $\mathcal{M}$ be a Lie group with bi-invariant metric. Set $\mathfrak{m} := T_I\mathcal{M}$. Let $(x_1(.), \ldots, x_N(.)) = \mathbf{x}(.) \in \mathcal{B}^N \subset \mathcal{M}^N$ and $(w_1(.), \ldots, w_N(.)) = \mathbf{w}(.) \in \mathfrak{m}^N$.

Prop. 1. Define **geodesic consensus error** as

$$\varphi(\mathbf{x}) := \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} d_g(x_i, x_j)^2$$

Then

$$\nabla\varphi(\mathbf{x}) = 0 \wedge \mathbf{x} \in \mathcal{B}^N \iff \varphi(\mathbf{x}) = 0 \iff x_1 = \cdots = x_N$$

Compactified dynamics:

$$\dot{\mathbf{x}} = -\nabla\varphi(\mathbf{x}) - dL_{\mathbf{x}}\mathbf{v} = 0 \quad (1)$$
$$\dot{\mathbf{w}} = \mathbf{L}\mathbf{v} = 0, \quad (2)$$

where $\mathbf{v} = -\mathbf{w} + \nabla\mathbf{f}(\mathbf{x})$, $\mathbf{f}(\mathbf{x}) := (f_1(x_1), \ldots, f_N(x_N))$, and $\mathbf{L} = L(G)$.

We don't want arbitrary consensus, we want consensus to $x^* = (\bar{z}, \ldots, \bar{z}) \in \mathcal{B}^N \subset \mathcal{M}^N$.

Proof overview continued

- **Prop. 2.** Bi-invariant metric implies $T^I_{x_i} \log_{x_i}(x_j) = -T^I_{x_j} \log_{x_j}(x_i) = \log\left(x_i^{-1} x_j\right)$.
- **Prop. 3.** $\sum_{i=1}^N w_i(t) = 0$ for all $t \geq 0$.
- **Prop. 4.** $\bar{z} = RCM(z_1, \ldots, z_N)$ iff $\sum_{i=1}^N \log\left(\bar{z}^{-1} z_i\right) = 0$.

Proof overview continued

- **Prop. 2.** Bi-invariant metric implies $T^I_{x_i} \log_{x_i}(x_j) = -T^I_{x_j} \log_{x_j}(x_i) = \log\left(x_i^{-1} x_j\right)$.
- **Prop. 3.** $\sum_{i=1}^N w_i(t) = 0$ for all $t \geq 0$.
- **Prop. 4.** $\bar{z} = RCM(z_1, \dots, z_N)$ iff $\sum_{i=1}^N \log\left(\bar{z}^{-1} z_i\right) = 0$.

Prop 1. and **Prob. 2.** implies $x_1 = \cdots = x_N =: x^*$ and $\mathbf{w} = \nabla \mathbf{f}(\mathbf{x}) = \left(\log\left((x^*)^{-1} z_i\right)\right)_i$.

Proof overview continued

- **Prop. 2.** Bi-invariant metric implies $T^I_{x_i} \log_{x_i}(x_j) = -T^I_{x_j} \log_{x_j}(x_i) = \log\left(x_i^{-1} x_j\right)$.
- **Prop. 3.** $\sum_{i=1}^N w_i(t) = 0$ for all $t \geq 0$.
- **Prop. 4.** $\bar{z} = RCM(z_1, \ldots, z_N)$ iff $\sum_{i=1}^N \log\left(\bar{z}^{-1} z_i\right) = 0$.

Prop 1. and **Prob. 2.** implies $x_1 = \cdots = x_N =: x^*$ and $\mathbf{w} = \nabla \mathbf{f}(\mathbf{x}) = (\log\left((x^*)^{-1} z_i\right))_i$.

Proof overview continued

- **Prop. 2.** Bi-invariant metric implies $T^I_{x_i} \log_{x_i}(x_j) = -T^I_{x_j} \log_{x_j}(x_i) = \log\left(x_i^{-1} x_j\right)$.
- **Prop. 3.** $\sum_{i=1}^N w_i(t) = 0$ for all $t \geq 0$.
- **Prop. 4.** $\bar{z} = RCM(z_1, \ldots, z_N)$ iff $\sum_{i=1}^N \log\left(\bar{z}^{-1} z_i\right) = 0$.

Prop 1. and **Prob. 2.** implies $x_1 = \cdots = x_N =: x^*$ and $\mathbf{w} = \nabla \mathbf{f}(\mathbf{x}) = (\log\left((x^*)^{-1} z_i\right))_i$.

Prop 3. and $\mathbf{w} = (\log\left((x^*)^{-1} z_i\right))_i$ imply $\sum_{i=1}^N \log\left((x^*)^{-1} z_i\right) = 0$

Proof overview continued

- **Prop. 2.** Bi-invariant metric implies $T^I_{x_i}\log_{x_i}(x_j) = -T^I_{x_j}\log_{x_j}(x_i) = \log\left(x_i^{-1}x_j\right)$.
- **Prop. 3.** $\sum_{i=1}^N w_i(t) = 0$ for all $t \geq 0$.
- **Prop. 4.** $\bar{z} = RCM(z_1, \ldots, z_N)$ iff $\sum_{i=1}^N \log\left(\bar{z}^{-1}z_i\right) = 0$.

Prop 1. and **Prob. 2.** implies $x_1 = \cdots = x_N =: x^*$ and $\mathbf{w} = \nabla\mathbf{f}(\mathbf{x}) = (\log\left((x^*)^{-1}z_i\right))_i$.

Prop 3. and $\mathbf{w} = (\log\left((x^*)^{-1}z_i\right))_i$ imply $\sum_{i=1}^N \log\left((x^*)^{-1}z_i\right) = 0$

Prop 4. and $\sum_{i=1}^N \log\left((x^*)^{-1}z_i\right) = 0$ implies $x^* = \bar{z}$

□

Comparisons

We compare against these algorithms:

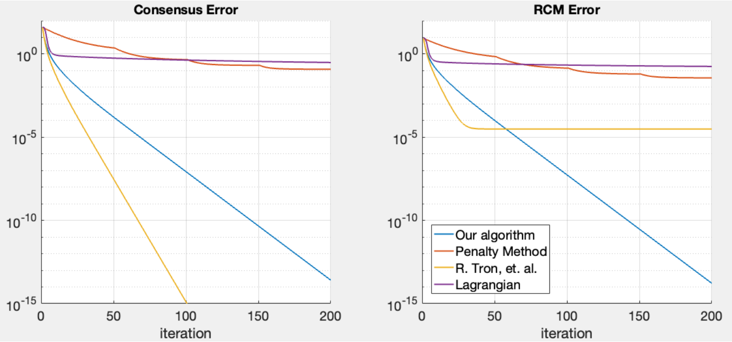
- (Lagrangian) Search for saddle points of the Lagrangian of the consensus reformulation:

$$\begin{cases} \min_{\mathbf{x}\in\mathcal{M}^N} f(\mathbf{x}) := \sum_{i=1}^N d_g^2(x_i, z_i) \\ \text{s.t. } \sum_{j\in\mathcal{N}_i} d_g(x_i, x_j)^2 = 0 \;\forall i \end{cases}$$

- (Penalty) Solve

$$\min_{x\in\mathcal{M}^N} \mu_k \sum_{i=1}^N d_g(x_i, z_i)^2 + \sum_{i=1}^N \sum_{j\in\mathcal{N}_i} d_g(x_i, x_j)^2$$

  via gradient descent for $k = 1, 2, \ldots$, where μ_k is an increasing sequence.



Concluding remarks

Summary:

- We generalized average consensus to Riemannian manifolds
- Convergence guarantees for Lie groups with bi-invariant metric
- Faster than the Lagrangian method empirically and has a seemingly linear rate of convergence

Future directions:

- Generalize to arbitrary Riemannian manifolds
- Investigate stationary points, including their dynamical and statistical properites
- Investigate relationship between the convergence rate and $G = ([N], E)$ or the curvature of $\mathcal{M}$