Two-sided Kirszbraun Theorem

Arturs Backurs
TTIC

Sepideh Mahabadi
TTIC

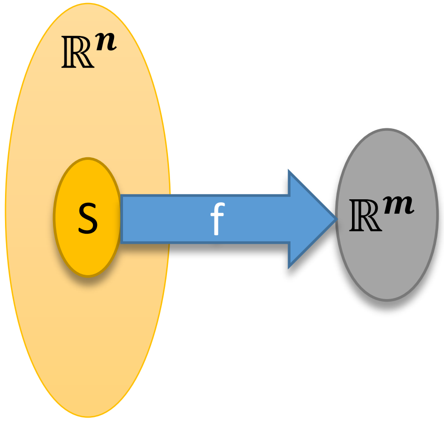
Konstantin Makarychev
Northwestern University

Yury Makarychev
TTIC

Plan

1. Background
   - Extension of functions
2. Our results
   - Two-sided Kirszbraun Theorem
3. Overview of the approach

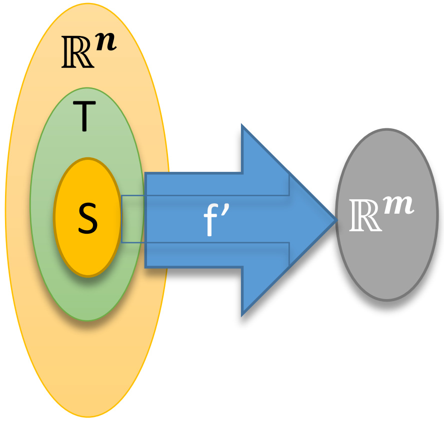
Extension of Functions



Notation throughout the talk

- We have a function $f: S \to \mathbb{R}^n$
- Which is defined over a subset $S \subset \mathbb{R}^m$

Extension of Functions

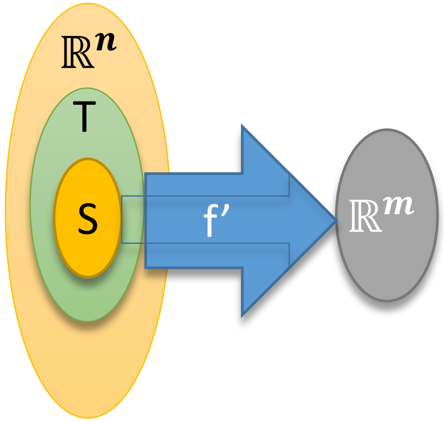


Notation throughout the talk

- We have a function $\boldsymbol{f}: \boldsymbol{S} \rightarrow \mathbb{R}^{\boldsymbol{n}}$
- Which is defined over a subset $\boldsymbol{S} \subset \mathbb{R}^{\boldsymbol{m}}$

Extensions of the map $\boldsymbol{f}$ to a superset $\boldsymbol{T}$ of $\boldsymbol{S}$

Extension of Functions



Notation throughout the talk

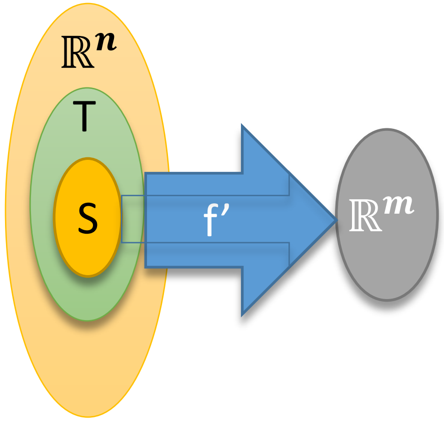
- We have a function $\boldsymbol{f}: \boldsymbol{S} \rightarrow \mathbb{R}^{\boldsymbol{n}}$
- Which is defined over a subset $\boldsymbol{S} \subset \mathbb{R}^{\boldsymbol{m}}$

> Extensions **of the map $\boldsymbol{f}$ to a superset S $\subset \boldsymbol{T} \subset \mathbb{R}^{\boldsymbol{n}}$**

❑ **Extension to the superset $\boldsymbol{T}$** , i.e., $\boldsymbol{f}': \boldsymbol{T} \rightarrow \mathbb{R}^{\boldsymbol{m}}$ so that

- $\boldsymbol{f}'(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$ for any $x \in S$
- Maintaining other properties …

Extension of Functions



Notation throughout the talk

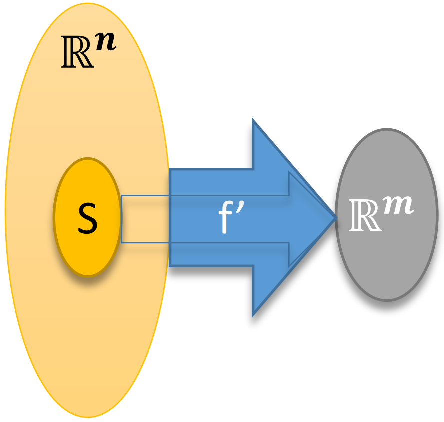
- We have a function $f: S \rightarrow \mathbb{R}^n$
- Which is defined over a subset $S \subset \mathbb{R}^m$

Extensions **of the map f to a superset $S \subset T \subset \mathbb{R}^n$**

- ❑ **Extension to the superset T**, i.e., $f': T \rightarrow \mathbb{R}^m$ so that
  - $f'(x) = f(x)$ for any $x \in S$
  - Maintaining other properties …

1. **Lipschitz Constant**
   (Lipschitz Extension)
2. **Bi-Lipschitz Constant, i.e., distortion**
   (Bi-Lipschitz Extension)

Extension of Functions



Notation throughout the talk

- We have a function $f: S \to \mathbb{R}^n$
- Which is defined over a subset $S \subset \mathbb{R}^m$

Extensions **of the map f to a superset T of A**

❑ **Extension to the whole $\mathbb{R}^n$**, i.e., $f': \mathbb{R}^n \to \mathbb{R}^m$ so that

- $f'(x) = f(x)$ for any $x \in A$
- Maintaining other properties …

1. **Lipschitz Constant**
   (Lipschitz Extension)
2. **Bi-Lipschitz Constant, i.e., distortion**
   (Bi-Lipschitz Extension)

Lipschitz Extension

- A map $f : X \to Y$ is **L-Lipschitz** if for all $x, x' \in X$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq \boldsymbol{L} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

→ Euclidean

Lipschitz Extension

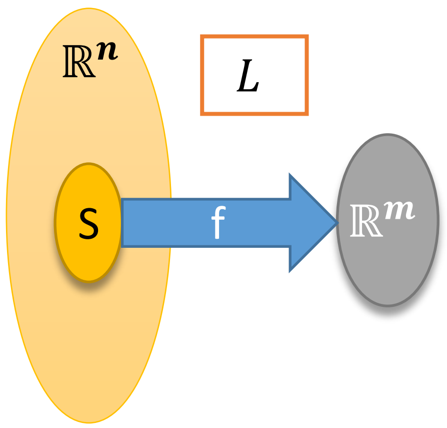
- A map $f: X \rightarrow Y$ is **L-Lipschitz** if for all $x, x' \in X$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq \boldsymbol{L} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

→ Euclidean

- **Lipschitz extension:**

  Given: a L-Lipschitz map $f: S \rightarrow \mathbb{R}^{\boldsymbol{m}}$, where $S \subseteq \mathbb{R}^{\boldsymbol{n}}$



Lipschitz Extension

- A map $f: X \to Y$ is ***L*-Lipschitz** if for all $x, x' \in X$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq \boldsymbol{L} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

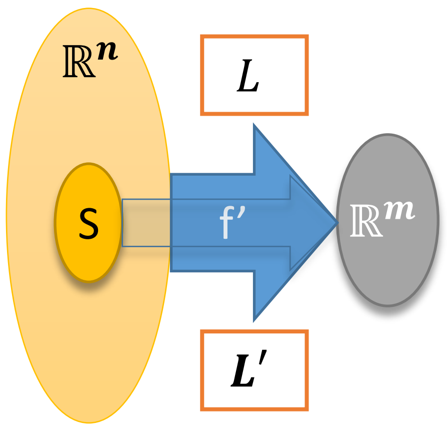
→ Euclidean

- **Lipschitz extension:**

  Given: a L-Lipschitz map $f: S \to \mathbb{R}^{\boldsymbol{m}}$, where $S \subseteq \mathbb{R}^{\boldsymbol{n}}$

  Goal: a map $f': \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ s.t.

  - f' is an extension of f
  - f' is L'-Lipschitz



Kirszbraun Extension Theorem

- A map $f: X \to Y$ is **L-Lipschitz** if for all $x, x' \in X$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq \boldsymbol{L} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

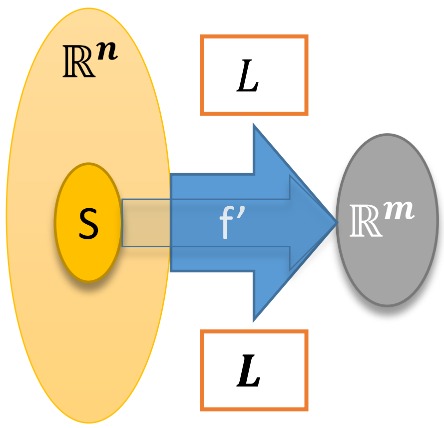
→ Euclidean

- **Lipschitz extension:**

  Given: a L-Lipschitz map $f: S \to \mathbb{R}^{\boldsymbol{m}}$, where $S \subseteq \mathbb{R}^{\boldsymbol{n}}$

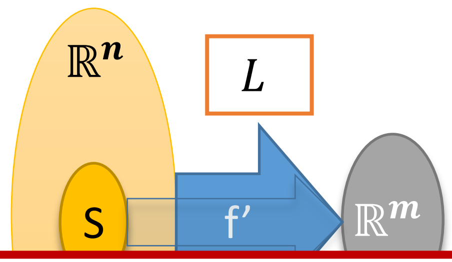
  Goal: a map $f': \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ s.t.

  - f' is an extension of f
  - f' is L'-Lipschitz



Kirszbraun extension theorem '34: for $\boldsymbol{S} \subset \mathbb{R}^{\boldsymbol{n}}$, every L-Lipschitz map $\boldsymbol{f}: \boldsymbol{S} \to \mathbb{R}^{\boldsymbol{m}}$ can be extended to the whole $\mathbb{R}^n$ keeping the **same** Lipschitz constant, i.e., $\boldsymbol{L}' = \boldsymbol{L}$.

Kirszbraun Extension Theorem



- A map $f: X \to Y$ is ***L*-Lipschitz** if for all $x, x' \in X$:

  $$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \le \boldsymbol{L} \cdot \|\boldsymbol{x}$$

- **Lipschitz extension:**

  Given: a L-Lipschitz map $f: S$

  Goal: a map $f': \mathbb{R}^n \to \mathbb{R}^m$ s.t.

  - f' is an extension of f
  - f' is L'-Lipschitz

- Applications
  - Prioritized and Terminal Dimension reduction
  - Clustering
  - …

Kirszbraun extension theorem '34: for $\boldsymbol{S} \subset \mathbb{R}^n$, every L-Lipschitz map $\boldsymbol{f}: \boldsymbol{S} \to \mathbb{R}^m$ can be extended to the whole $\mathbb{R}^n$ keeping the **same** Lipschitz constant, i.e., $\boldsymbol{L}' = \boldsymbol{L}$.

f' can decrease distances arbitrarily

f' can decrease distances arbitrarily

- Can we get any lower bound?

f' can decrease distances arbitrarily

- Can we get any lower bound?
  - **Bi-Lipschitz extension: [MMMR'18]**
    - Initial map $f: X \to Y$ is **D-bi-Lipschitz** or has **distortion D**, i.e., for some λ and all $x, x' \in X$:

$$\lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq D \cdot \lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

f' can decrease distances arbitrarily

- Can we get any lower bound?
  - **Bi-Lipschitz extension: [MMMR'18]**
    - Initial map $f: X \to Y$ is **D-bi-Lipschitz** or has **distortion D**, i.e., for some λ and all $x, x' \in X$:
      $$\lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq D \cdot \lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$
    - f' will have distortion $O(D)$
    - Using "extra coordinates"

f' can decrease distances arbitrarily

- Can we get any lower bound?
  - **Bi-Lipschitz extension: [MMMR'18]**
    - Initial map $f: X \to Y$ is **D-bi-Lipschitz** or has **distortion D**, i.e., for some λ and all $x, x' \in X$:
    $$\lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq D \cdot \lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$
    - f' will have distortion $O(D)$
    - Using "extra coordinates"

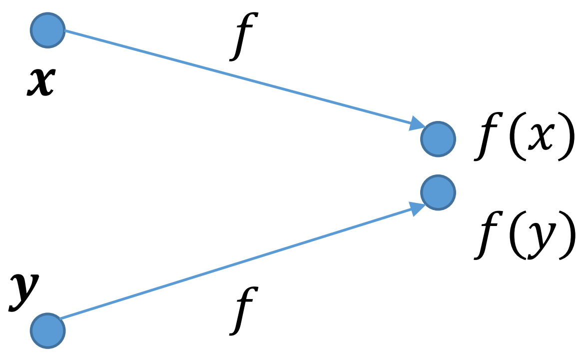
➢ **What if we have no such guarantee?**

Two-Sided Kirszbraun Theorem

- Decreasing distances is unavoidable

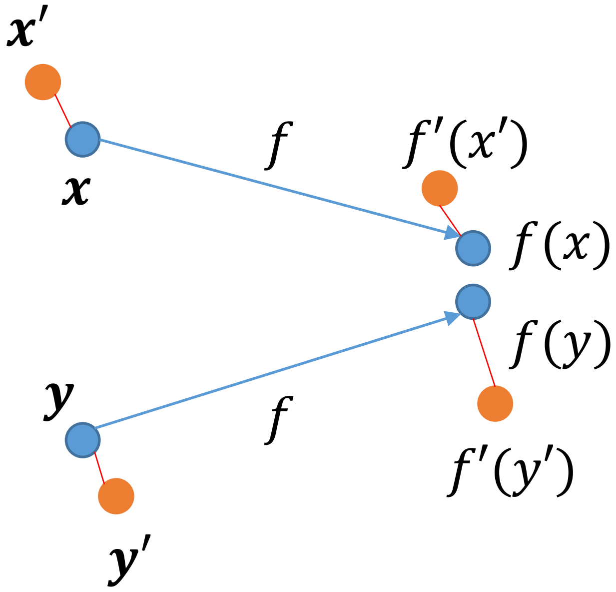
Two-Sided Kirszbraun Theorem

- Decreasing distances is unavoidable
  - $x, y \in S$
  - $\|f(x) - f(y)\|$ is small



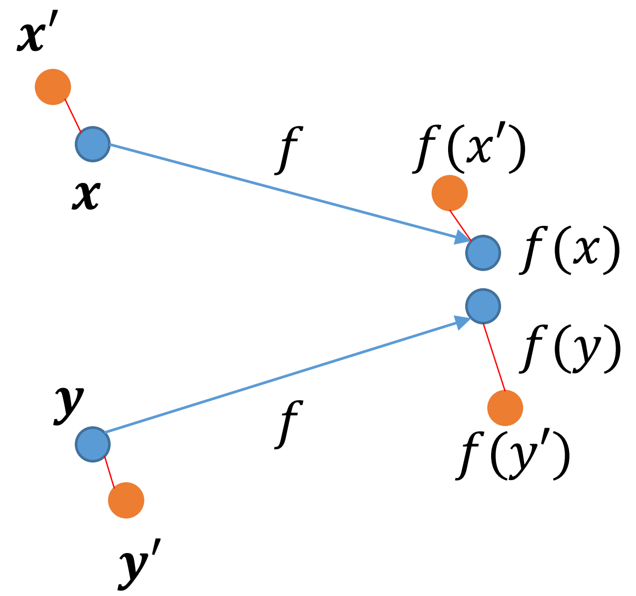
Two-Sided Kirszbraun Theorem

- Decreasing distances is unavoidable
  - $x, y \in S$
  - $\|f(x) - f(y)\|$ is small
  - $x' \in T$ is close to x, and $y' \in T$ is close to y



Two-Sided Kirszbraun Theorem

- ❑ Decreasing distances is unavoidable
  - $x, y \in S$
  - $\|f(x) - f(y)\|$ is small
  - $x' \in T$ is close to x, and $y' \in T$ is close to y



➢ **Question:** Can we decrease distances between any pair of points as little as possible?

Plan

1. Background
   - Extension of functions
2. Our results
   - Two-sided Kirszbraun Theorem
3. Overview of the approach

Results in a nutshell

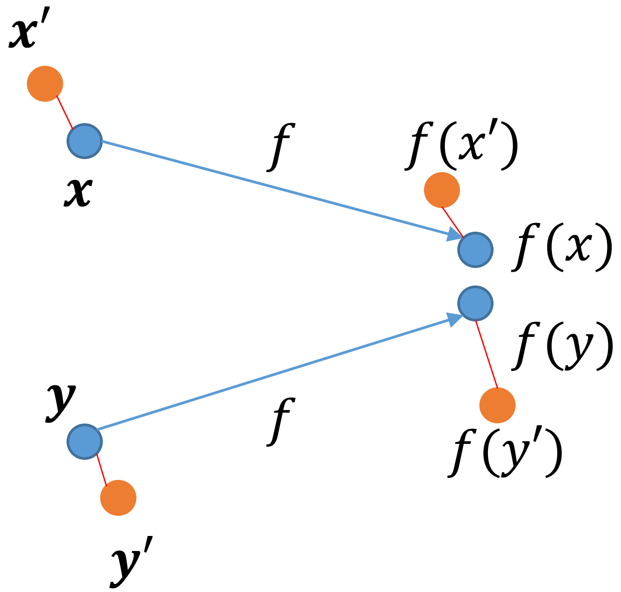
A “tight” variant of the Kirszbraun theorem:

It is possible to find an extension map f' such that the distance between any pair of points is not decreased by more than what is “necessary”.

What is necessary? An upper bound

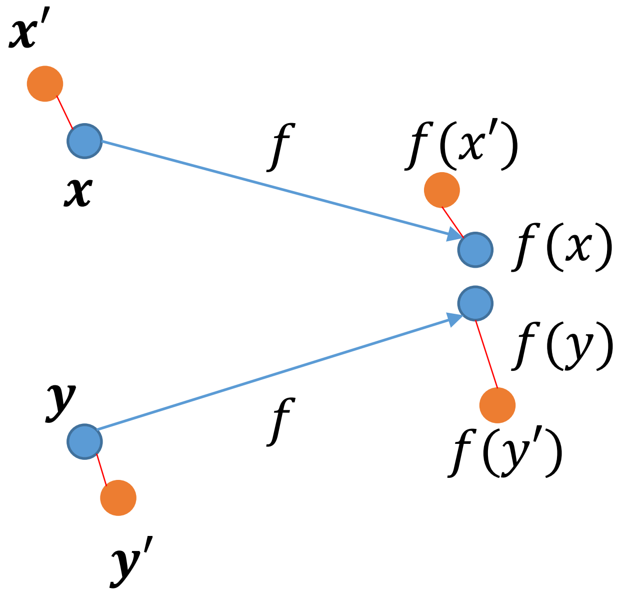
What is necessary? An upper bound

$$(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\|)$$



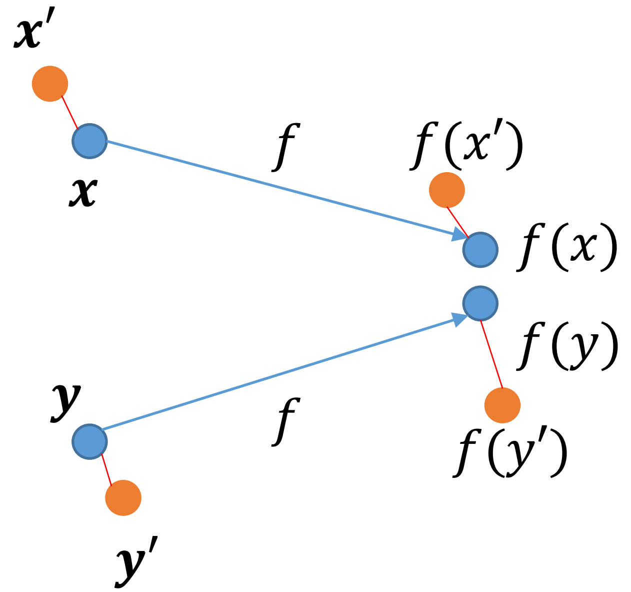
What is necessary? An upper bound

$$\inf_{x,y \in S} \left(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\| \right)$$



What is necessary? An upper bound

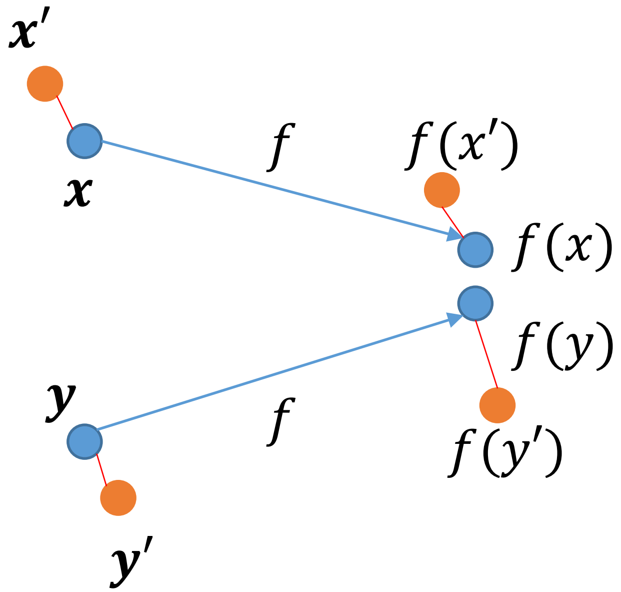
$$\min(L\|x' - y'\|, \inf_{x,y \in S}(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\|)$$



What is necessary? An upper bound

- Define metric

$$\boldsymbol{d_{ub}}(\boldsymbol{x'}, \boldsymbol{y'}) = \min(L\|x' - y'\|, \inf_{x,y \in S}(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\|)$$

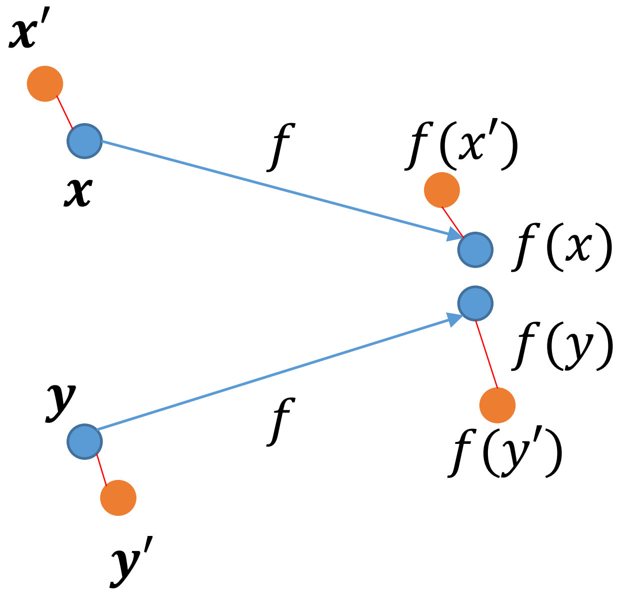


What is necessary? An upper bound

- Define metric

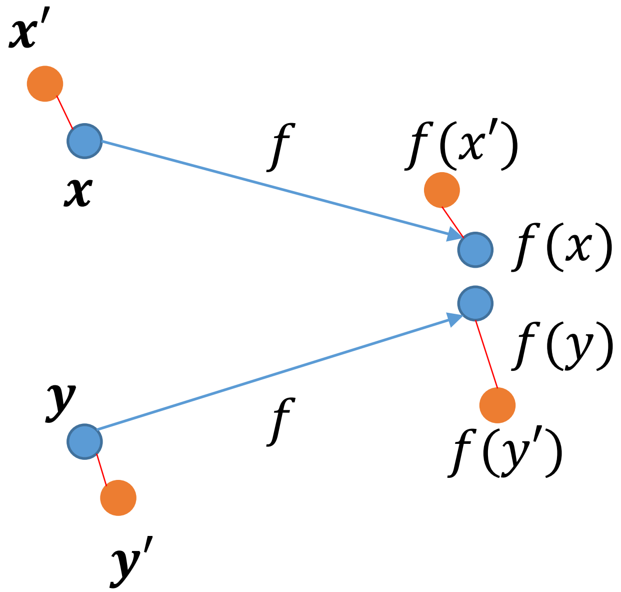
$$\boldsymbol{d_{ub}}(\boldsymbol{x}', \boldsymbol{y}') = \min(L\|x' - y'\|, \inf_{x,y \in S}(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\|)$$

- cL-Lipschitz extension $f': \mathrm{T} \to \mathbb{R}^m$ has $\|f'(x') - f'(y')\| \leq c d_{ub}(x', y')$



What is necessary? An upper bound

- Define metric
$$\boldsymbol{d_{ub}}(\boldsymbol{x'}, \boldsymbol{y'}) = \min\left(L\|x' - y'\|, \inf_{x,y \in S}\left(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\|\right)\right)$$
- cL-Lipschitz extension $f' \colon \mathrm{T} \to \mathbb{R}^m$ has $\|f'(x') - f'(y')\| \leq c\, d_{ub}(x', y')$
- **Can we find $\boldsymbol{f'}$ such that $\|\boldsymbol{f'}(\boldsymbol{x'}) - \boldsymbol{f'}(\boldsymbol{y'})\| \geq \boldsymbol{\Omega}(\boldsymbol{d_{ub}}(\boldsymbol{x'}, \boldsymbol{y'}))$?**

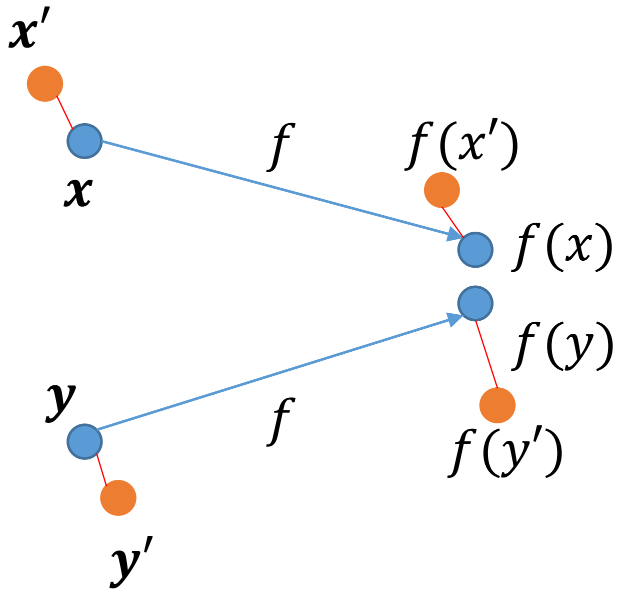


What is necessary? An upper bound

- Define metric

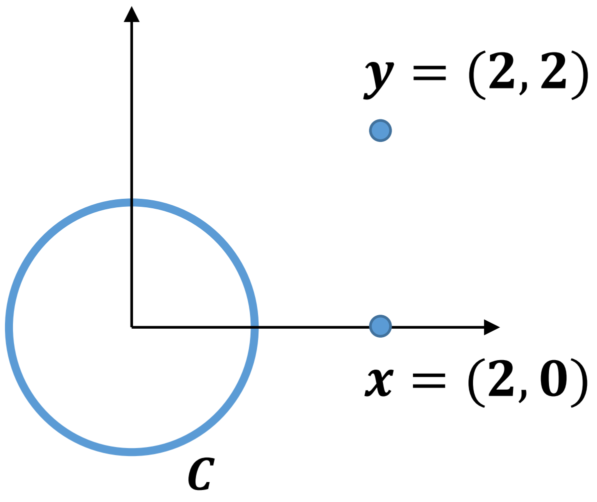
$$\boldsymbol{d_{ub}}(\boldsymbol{x'}, \boldsymbol{y'}) = \min(L\|x' - y'\|, \inf_{x,y \in S}(L\|x' - x\| + \|f(x) - f(y)\| + L\|y - y'\|))$$

- cL-Lipschitz extension $f': \mathrm{T} \to \mathbb{R}^m$ has $\|f'(x') - f'(y')\| \leq c d_{ub}(x', y')$
- **Can we find $\boldsymbol{f'}$ such that $\|\boldsymbol{f'}(\boldsymbol{x'}) - \boldsymbol{f'}(\boldsymbol{y'})\| \geq \boldsymbol{\Omega}(\boldsymbol{d_{ub}}(\boldsymbol{x'}, \boldsymbol{y'}))$?**
    - Short Answer: No
    - Long Answer: We need extra relaxations



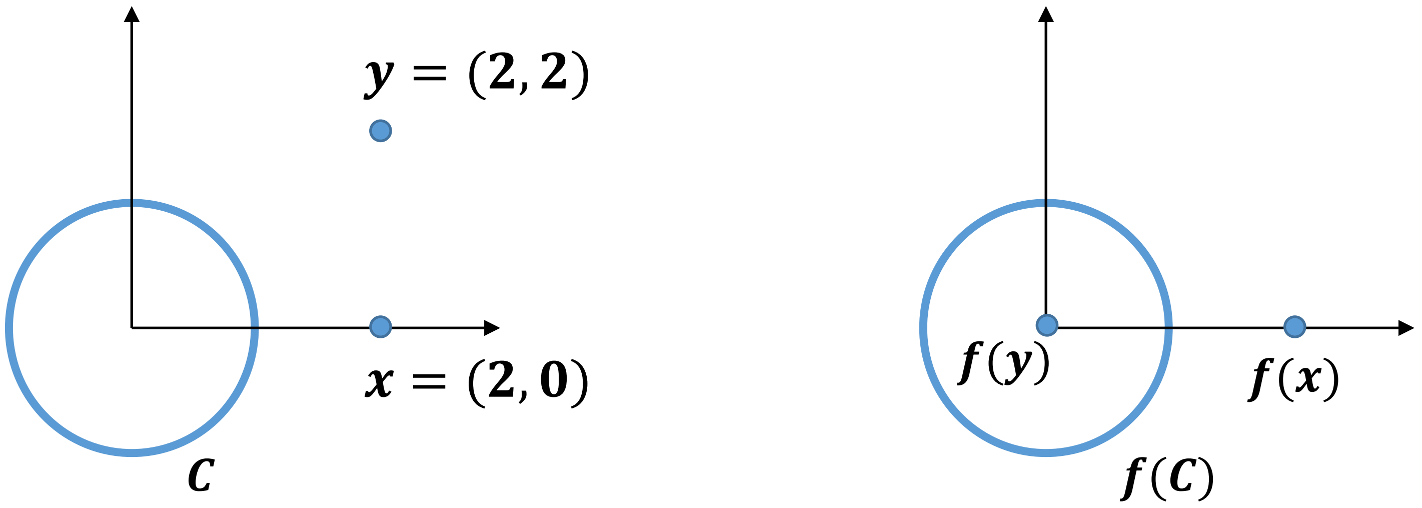
A bad example

- $S = C \cup \{x, y\}$



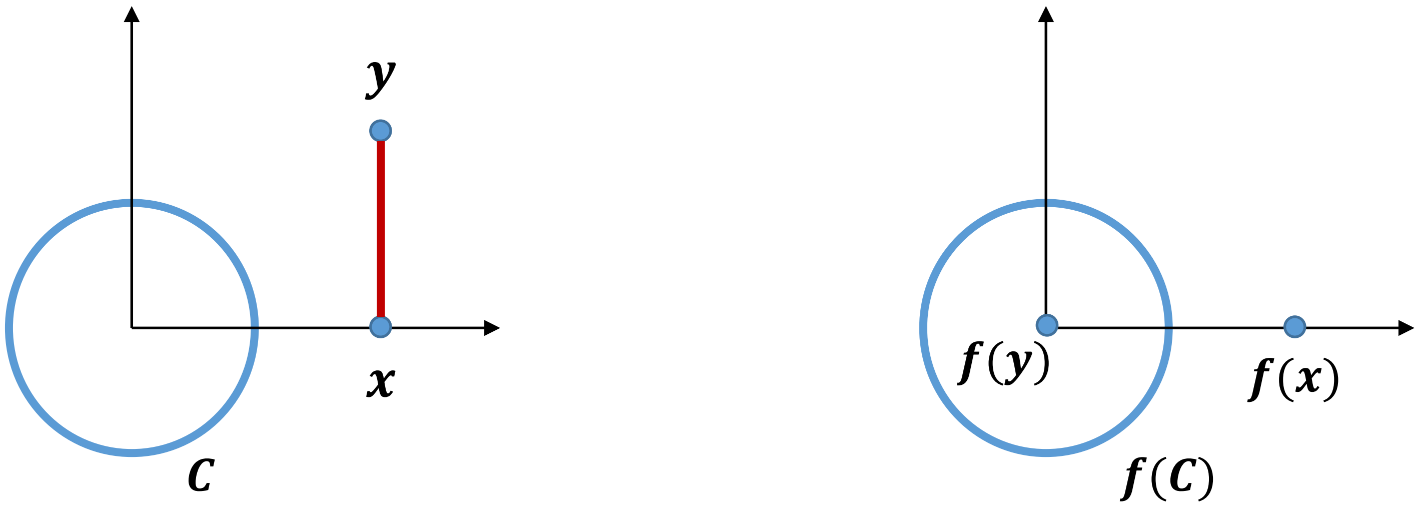
A bad example

- $S = C \cup \{x, y\}$
- f is 1-Lipschitz



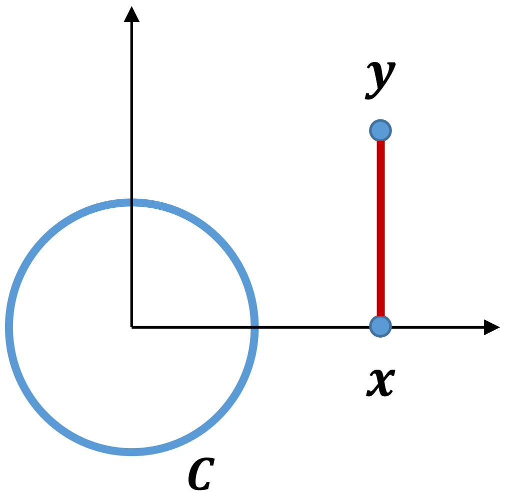
A bad example

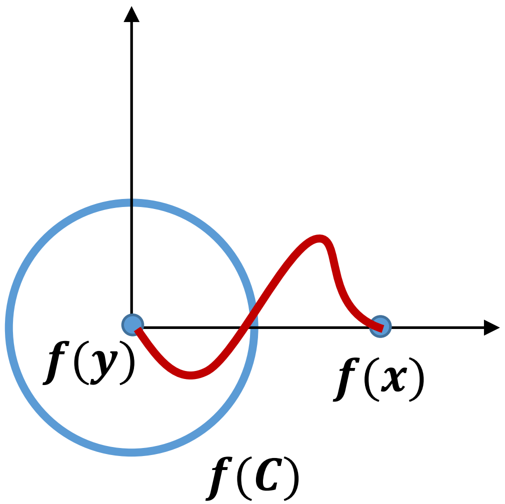
- $S = C \cup \{x, y\}$, $T = C \cup [x, y]$
- f is 1-Lipschitz



A bad example

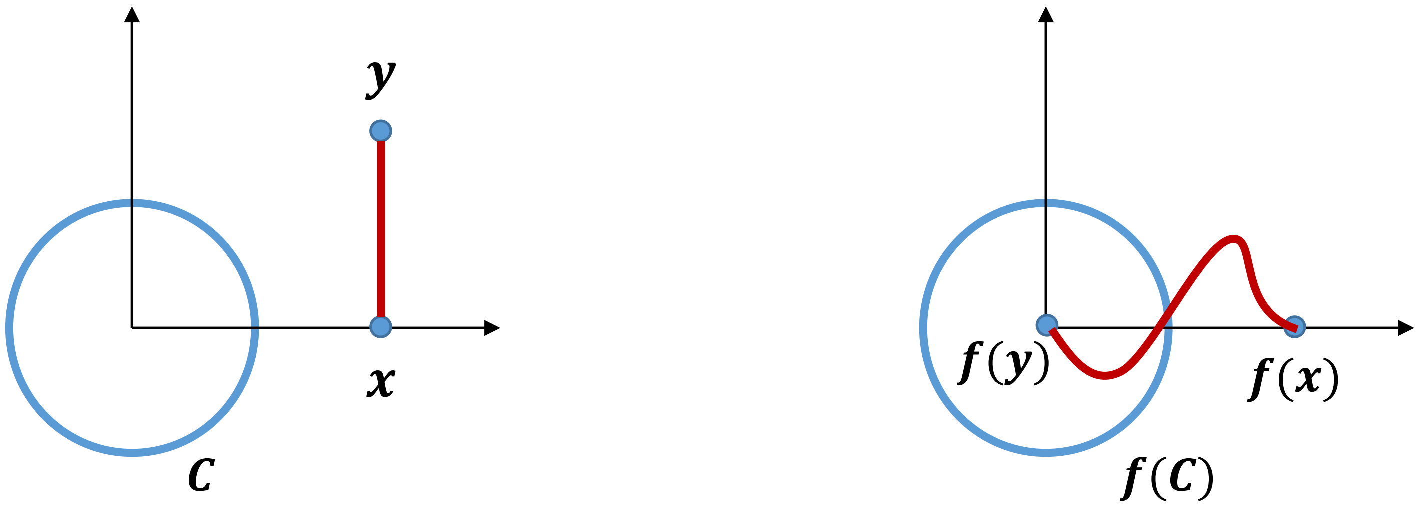
- $S = C \cup \{x, y\}, T = C \cup [x, y]$
- f is 1-Lipschitz
- $d_{ub} > 0$ for every pair of points $u, v \in T$
- But the image of $[x, y]$ intersects the circle, i.e., $f'(u) - f'(v) = 0$





A bad example

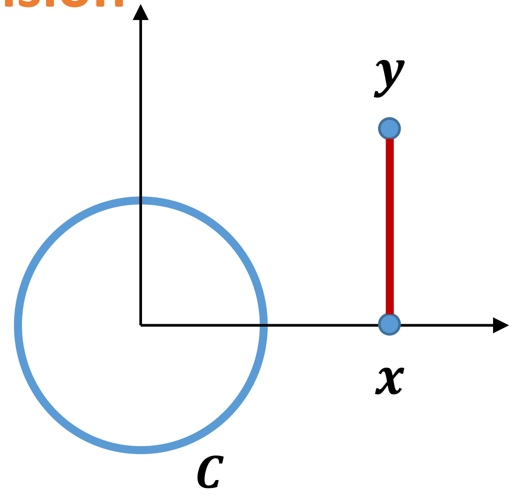
- $S = C \cup \{x, y\}, T = C \cup [x, y]$
- f is 1-Lipschitz
- $d_{ub} > 0$ for every pair of points $u, v \in T$
- But the image of $[x, y]$ intersects the circle, i.e., $f'(u) - f'(v) = 0$, so it is not possible to get $\|f'(x') - f'(y')\| \geq \Omega(d_{ub}(x', y'))$

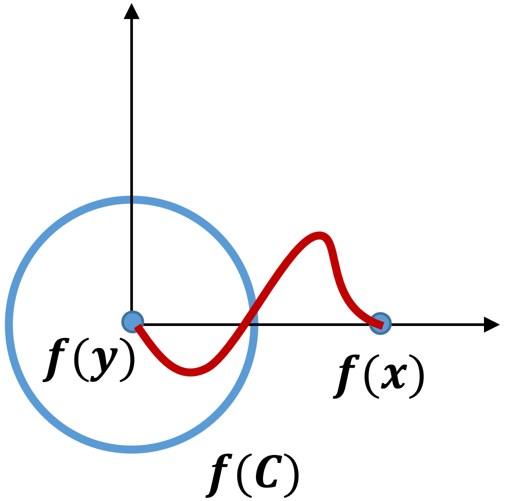


A bad example

- $S = C \cup \{x, y\}, T = C \cup [x, y]$
- f is 1-Lipschitz
- $d_{ub} > 0$ for every pair of points $u, v \in T$
- But the image of $[x, y]$ intersects the circle, i.e., $f'(u) - f'(v) = 0$, so it is not possible to get $\|f'(x') - f'(y')\| \geq \Omega(d_{ub}(x', y'))$

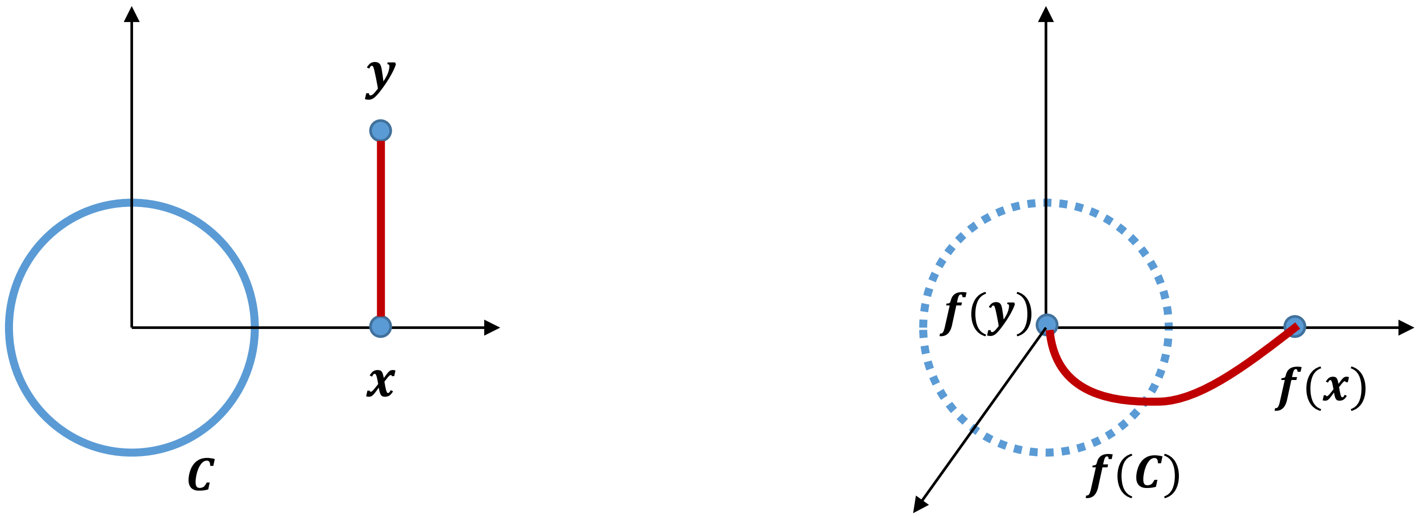
➢ **Outer extension**





Relaxation I: Outer Extension

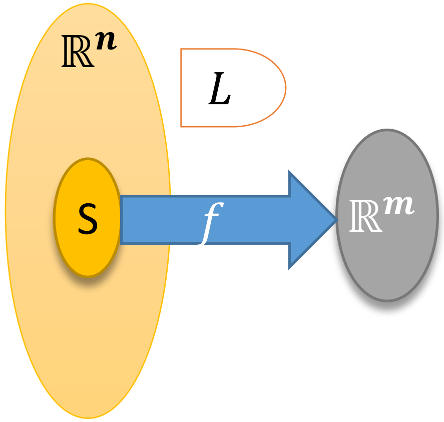
- Use additional coordinates in the image of the extended map



Lipschitz Outer-Extension

Given: a map $f: S \to \mathbb{R}^m$, where

- $S \subseteq T \subset \mathbb{R}^n$
- f is L −Lipschitz



Lipschitz Outer-Extension

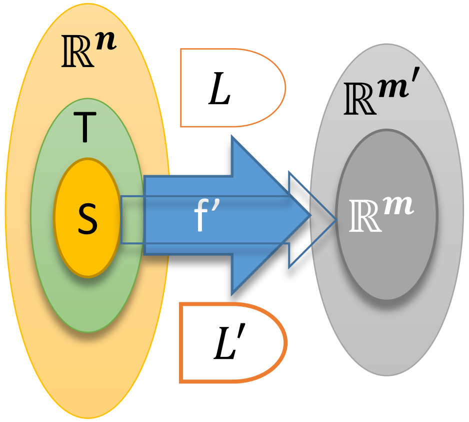
Given: a map $f: S \to \mathbb{R}^m$, where

- $S \subseteq T \subset \mathbb{R}^n$
- f is L –Lipschitz

Goal: a map $f': X \to \mathbb{R}^{m'}$, where

- $m' > m$
- f' is L' –Lipschitz
- f' is an **(outer)-extension** of f: for every $x \in S$

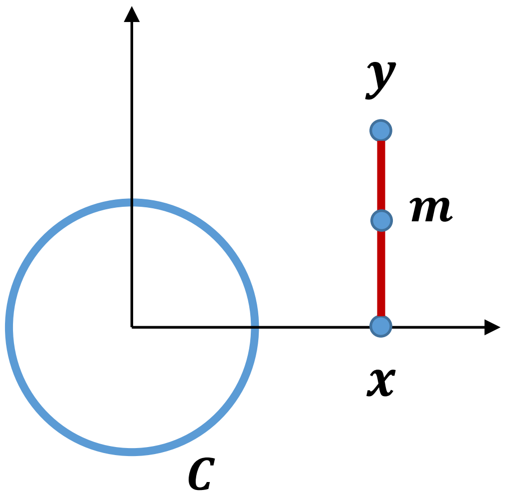
$$f'(x) = f(x) \oplus \underbrace{(0, \dots, 0)}_{m' - m}$$

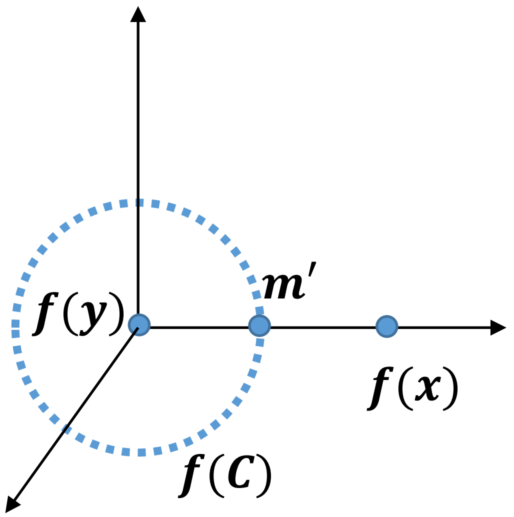


Relaxation II: increase the Lipschitz constant

Relaxation II: increase the Lipschitz constant

- If f' must have Lipschitz constant equal to 1, then $\boldsymbol{m}$ should be mapped to $\boldsymbol{m}'$
  - the distances would decrease infinitely.

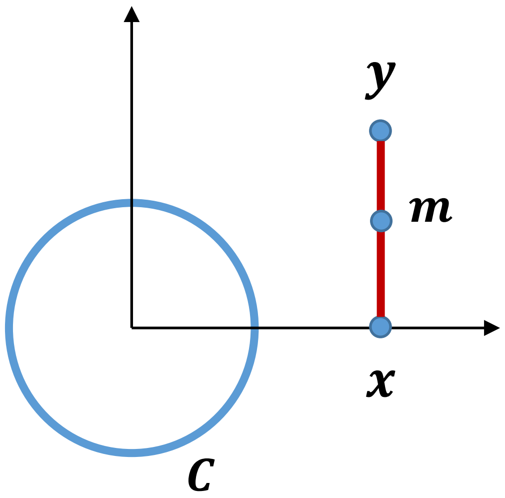


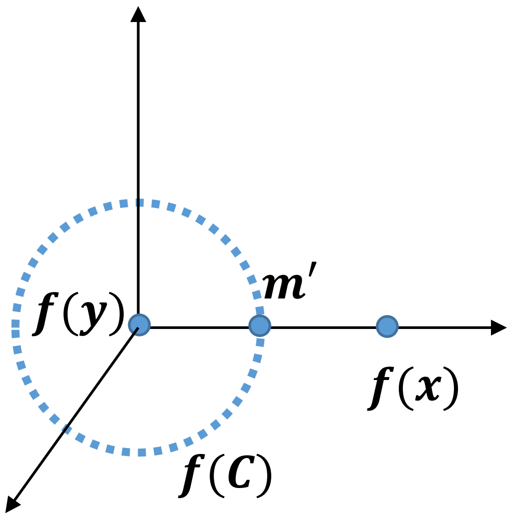


Relaxation II: increase the Lipschitz constant

- If f' must have Lipschitz constant equal to 1, then $\boldsymbol{m}$ should be mapped to $\boldsymbol{m}'$
  - the distances would decrease infinitely.

➢ Instead find a **$(1+\epsilon)L$-extension** f'





Results: Two-sided Kirszbraun Theorem

Given:

- $f: S \to \mathbb{R}^m$ is L −Lipschitz
- $S \subset T \subset \mathbb{R}^n$

Find: the extended map $f': T \to \mathbb{R}^m \oplus \mathbb{R}^\Delta \approx \mathbb{R}^{m'}$ such that

- f' is $(1+\epsilon)L$ −Lipschitz
- $\|f'(x) - f'(y)\| \geq c\sqrt{\epsilon} d_{ub}(x, y)$ for all $x, y \in T$
- If $|T \setminus S|$ is finite, then $\Delta = O(\log |T \setminus S|)$.
- Otherwise $\Delta = \infty$

Pros

- Least Possible contraction: for any pair simultaneously upto a factor of $O\left(\sqrt{\epsilon}\right)$, i.e., Bound $\frac{\|f'(x)-f'(y)\|}{d_{ub}(x,y)} \in [c\sqrt{\epsilon}, 1+\epsilon]$

Pros

- Least Possible contraction: for any pair simultaneously upto a factor of $O(\sqrt{\epsilon})$, i.e., Bound $\frac{\|f'(x)-f'(y)\|}{d_{ub}(x,y)} \in [c\sqrt{\epsilon}, 1+\epsilon]$
- Easy to compute distances $||f'(x) - f'(y)||$
  - Computing $||f'(x) - f'(y)||$ in Kirszbraun theorem requires computing the entire map itself, which can be done using SDP
  - Here, we can just compute use $d_{ub}(x, y)$ as a good approximation

Pros

- **Least Possible contraction**: for any pair simultaneously upto a factor of $O(\sqrt{\epsilon})$, i.e., Bound $\frac{\|f'(x)-f'(y)\|}{d_{ub}(x,y)} \in [c\sqrt{\epsilon}, 1+\epsilon]$
- **Easy to compute distances** $\|f'(x) - f'(y)\|$
  - Computing $\|f'(x) - f'(y)\|$ in Kirszbraun theorem requires computing the entire map itself, which can be done using SDP
  - Here, we can just compute use $d_{ub}(x,y)$ as a good approximation
- **Optimal Parameters** (See next slide)

Lower bound results

1. $\sqrt{\epsilon}$ loss is required: There exists S and $T = S \cup \{z_1, z_2\}$ and a 1-Lipschitz function f s.t. for any $(1+\epsilon)$-Lipschitz extension of f, their distance has to decrease by a factor of $\sqrt{\epsilon}$, i.e., $\|f'(z_1) - f'(z_2)\| \leq O(\sqrt{\epsilon} d_{ub}(z_1, z_2))$

Lower bound results

1. $\sqrt{\epsilon}$ loss is required: There exists S and $T = S \cup \{z_1, z_2\}$ and a 1-Lipschitz function f s.t. for any $(1+\epsilon)$-Lipschitz extension of f, their distance has to decrease by a factor of $\sqrt{\epsilon}$, i.e., $\|f'(z_1) - f'(z_2)\| \leq O(\sqrt{\epsilon} d_{ub}(z_1, z_2))$
2. $\log |T \setminus S|$ dimensions is required for finite sets: for any m, n, N, there exists an instance s.t. $|T \setminus S| = N$, and any outer Lipschitz extension with $\|f'(x) - f'(y)\| \geq \boldsymbol{c} d_{ub}(x, y)$ requires $m' = c' \log N$ where $c' = 1/\log(\frac{L}{c} + 1)$

Lower bound results

1. $\sqrt{\epsilon}$ loss is required: There exists S and $T = S \cup \{z_1, z_2\}$ and a 1-Lipschitz function f s.t. for any $(1+\epsilon)$-Lipschitz extension of f, their distance has to decrease by a factor of $\sqrt{\epsilon}$, i.e., $\|f'(z_1) - f'(z_2)\| \leq O(\sqrt{\epsilon} d_{ub}(z_1, z_2))$
2. $\log |T \setminus S|$ dimensions is required for finite sets: for any m, n, N, there exists an instance s.t. $|T \setminus S| = N$, and any outer Lipschitz extension with $\|f'(x) - f'(y)\| \geq c d_{ub}(x, y)$ requires $m' = c' \log N$ where $c' = 1/\log(\frac{L}{c} + 1)$
3. Infinite dimension is required for infinite sets: for any m, n, there exists an instance with infinite sets $S \subset T$, s.t. any outer Lipschitz extension with $\|f'(x) - f'(y)\| \geq c d_{ub}(x, y)$ for some c, requires $m' = \infty$

Application I: Bi-Lipschitz extension

- Our results immediately implies bi-Lipschitz extension of [MMakarychevMakarychevRazenshteyn’18]
- $O(D)$ distortion
- Caveat: we don’t have $m' = m + n$,
- Pros: easy to compute distances approximately.

Application II: Updating Euclidean Metric

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.
- We get additional information on a subset $\boldsymbol{Y} \subset X$ of them and compute the new Euclidean distance $\boldsymbol{d_Y}$

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.
- We get additional information on a subset $\boldsymbol{Y} \subset X$ of them and compute the new Euclidean distance $\boldsymbol{d_Y}$
- **Goal:** Combine into a distance function $\boldsymbol{d_c(x,y)} = \begin{cases} d_Y(x,y), & x,y \in Y \\ d_X(x,y), & otherwise \end{cases}$

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.
- We get additional information on a subset $\boldsymbol{Y} \subset X$ of them and compute the new Euclidean distance $\boldsymbol{d_Y}$
- **Goal:** Combine into a distance function $\boldsymbol{d_c(x,y)} = \begin{cases} d_Y(x,y), & x,y \in Y \\ d_X(x,y), & otherwise \end{cases}$
- Not necessarily a metric; Compute shortest path on it be $\boldsymbol{d_u}$

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.
- We get additional information on a subset $\boldsymbol{Y} \subset X$ of them and compute the new Euclidean distance $\boldsymbol{d_Y}$
- **Goal:** Combine into a distance function $\boldsymbol{d_c(x,y)} = \begin{cases} d_Y(x,y), & x,y \in Y \\ d_X(x,y), & otherwise \end{cases}$
- Not necessarily a metric; Compute shortest path on it be $\boldsymbol{d_u}$
- We want d_u to be Euclidean but it does not have to be

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.
- We get additional information on a subset $\boldsymbol{Y} \subset X$ of them and compute the new Euclidean distance $\boldsymbol{d_Y}$
- **Goal:** Combine into a distance function $\boldsymbol{d_c(x,y)} = \begin{cases} d_Y(x,y), & x,y \in Y \\ d_X(x,y), & otherwise \end{cases}$
- Not necessarily a metric; Compute shortest path on it be $\boldsymbol{d_u}$
- We want d_u to be Euclidean but it does not have to be

Theorem:

- Sufficient conditions for it: if $\boldsymbol{d_Y(x,y) \leq C d_X(x,Y)}$ for all $\boldsymbol{x,y \in Y}$, then the updated metric is $\boldsymbol{O(CAB)}$-Euclidean, where we assume d_x is A −Euclidean and d_y is B −Euclidean

Application II: Updating Euclidean Metric

- Dataset of objects $\boldsymbol{X}$ with a metric $\boldsymbol{d_X}$ resembling their similarity.
- We get additional information on a subset $\boldsymbol{Y} \subset X$ of them and compute the new Euclidean distance $\boldsymbol{d_Y}$
- **Goal:** Combine into a distance function $\boldsymbol{d_c(x,y)} = \begin{cases} d_Y(x,y), & x,y \in Y \\ d_X(x,y), & otherwise \end{cases}$
- Not necessarily a metric; Compute shortest path on it be $\boldsymbol{d_u}$
- We want d_u to be Euclidean but it does not have to be

Theorem:

- Sufficient conditions for it: if $\boldsymbol{d_Y(x,y) \leq C d_X(x,Y)}$ for all $\boldsymbol{x,y \in Y}$, then the updated metric is $\boldsymbol{O(CAB)}$-Euclidean, where we assume d_x is A −Euclidean and d_y is B −Euclidean
- Lower bound: The above condition is necessary otherwise one gets at least $\Omega(\log N)$ distiortion

Plan

1. Background
    - Extension of functions
2. Our results
    - Two-sided Kirszbraun Theorem
3. Overview of the approach

Overall approach

- $\boldsymbol{f}'(\boldsymbol{x}) = \boldsymbol{g}'(\boldsymbol{x}) \oplus \boldsymbol{h}'(\boldsymbol{x})$

Overall approach

- $\boldsymbol{f'(x) = g'(x) \oplus h'(x)}$

1. $g'(x)$: Kirszbraun extension of $f(x)$.
    - Does not increase distances
    - The same as $f(x)$ on the points in S

Overall approach

- $\boldsymbol{f'(x) = g'(x) \oplus h'(x)}$

1. $g'(x)$: Kirszbraun extension of $f(x)$.
   - Does not increase distances
   - The same as $f(x)$ on the points in S
2. $h'(x) = c\sqrt{\epsilon}L\, h(x)$:
   - $h(x)$ should be 0 when $x \in S$
   - Increases as a function of $R_x \coloneqq dist(x, S)$
   - $\|h(x) - h(y)\| \approx \Theta(\min(\|x - y\|, R_x + R_y))$
   - Use [Mendel&Naor'04] embedding (rescaled and truncated)

Construction of h

Two ingredients:

1. **[Mendel&Naor'04]:**
   - For any $r > 0$, there exists a map ψ_r from ℓ_2^n to the infinite dimensional sphere of radius r, such that it approximately preserve distances of value at most r.
   - $\|\psi(x) - \psi(y)\| = \Theta(\min\left(\|x - y\|, \sqrt{2}r\right))$

Construction of h

Two ingredients:

1. **[Mendel&Naor'04]:**
   - For any $r > 0$, there exists a map ψ_r from ℓ_2^n to the infinite dimensional sphere of radius r, such that it approximately preserves distances of value at most r.
   - $\|\psi(x) - \psi(y)\| = \Theta(\min(\|x - y\|, \sqrt{2}r))$

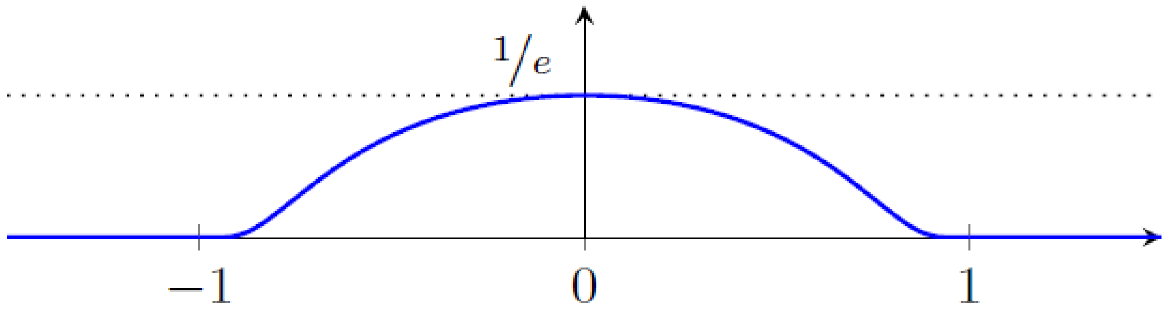
Set $r \approx R_x$

Construction of h

Two ingredients:

1. **[Mendel&Naor'04]:**
   - For any $r > 0$, there exists a map ψ_r from ℓ_2^n to the infinite dimensional sphere of radius r, such that it approximately preserves distances of value at most r.
   - $\|\psi(x) - \psi(y)\| = \Theta(\min\left(\|x-y\|, \sqrt{2}r\right))$
2. **Bump Function:**

$$\lambda(t) = \begin{cases} e^{-\frac{1}{1-t^2}}, & \text{if } t \in (-1,1) \\ 0, & \text{otherwise} \end{cases}$$



Overall approach

- $f'(x) = g'(x) \oplus h'(x)$

1. $g'(x)$: Kirszbraun extension of $f(x)$.
   - Does not increase distances
   - The same as $f(x)$ on the points in S
2. $h'(x) = c\sqrt{\epsilon}L\, h(x)$:
   - $h(x)$ should be 0 when $x \in S$
   - Increases as a function of $R_x \coloneqq dist(x, S)$
   - $\|h(x) - h(y)\| \approx \Theta(\min(\|x - y\|, R_x + R_y))$
   - Use [Mendel&Naor'04] embedding (rescaled and truncated)
   - For finite set $|T \setminus S|$, **we apply JL on top of $h'(x)$** to get the desired bound on the dimension

Summary

- Showed two sided variant of the Kirszbraun theorem
- It achieves asymptotically optimal parameters.
- Provides a simple approximate formula for computing distances
- Applications of our results to bi-Lip extension & Updating Euclidean metric.

Given:

- $f: S \to \mathbb{R}^m$ is L −Lipschitz
- $S \subset T \subset \mathbb{R}^n$

Find: the extended map $f': T \to \mathbb{R}^m \oplus \mathbb{R}^\Delta \approx \mathbb{R}^{m'}$ such that

- f' is $(1+\epsilon)L$ −Lipschitz
- $\|f'(x) - f'(y)\| \geq c\sqrt{\epsilon} d_{ub}(x, y)$ for all $x, y \in T$
- If $|T \setminus S|$ is finite, then $\Delta = O(\log |T \setminus S|)$.
- Otherwise $\Delta = \infty$

Summary

Thanks! Questions?

- Showed two sided variant of the Kirszbraun theorem
- It achieves asymptotically optimal parameters.
- Provides a simple approximate formula for computing distances
- Applications of our results to Updating Euclidean metric.

Given:

- $f: S \to \mathbb{R}^m$ is L −Lipschitz
- $S \subset T \subset \mathbb{R}^n$

Find: the extended map $f': T \to \mathbb{R}^m \oplus \mathbb{R}^\Delta \approx \mathbb{R}^{m'}$ such that

- f' is $(1+\epsilon)L$ −Lipschitz
- $\|f'(x) - f'(y)\| \geq c\sqrt{\epsilon} d_{ub}(x, y)$ for all $x, y \in T$
- If $|T \setminus S|$ is finite, then $\Delta = O(\log |T \setminus S|)$.
- Otherwise $\Delta = \infty$