On the Estimation of Distances Using Graph Distances

Ery Arias-Castro

University of California, San Diego

Collaborators

- Clément Berenfeld (École Normale Supérieure)
- Antoine Channarond (Université de Rouen)
- Thibaut Le Gouic (École Centrale de Marseille)
- Adel Javanmard (University of Southern California)
- Bruno Pelletier (Université de Rennes II)
- Nicolas Verzelen (INRA, Montpellier)

Euclidean distance matrix completion

Undirected graph: $G = (V, E)$ where $V = \{1, \ldots, n\}$.

δ_{ij} = dissimilarity between i and j, available when $(i, j) \in E$.

Euclidean distance matrix completion

Undirected graph: $G = (V, E)$ where $V = \{1, \ldots, n\}$.

δ_{ij} = dissimilarity between i and j, available when $(i, j) \in E$.

$D = (d_{ij}) \in \mathbb{R}^{n \times n}$ is a Euclidean distance matrix (EDM) if there is $y_1, \ldots, y_n$ in some Euclidean space such that

$$d_{ij} = \|y_i - y_j\|, \quad \forall i, j \in V$$

Euclidean distance matrix completion

Undirected graph: $G = (V, E)$ where $V = \{1, \dots, n\}$.

δ_{ij} = dissimilarity between i and j, available when $(i, j) \in E$.

$D = (d_{ij}) \in \mathbb{R}^{n \times n}$ is a Euclidean distance matrix (EDM) if there is $y_1, \dots, y_n$ in some Euclidean space such that

$$d_{ij} = \|y_i - y_j\|, \quad \forall i, j \in V$$

Euclidean distance matrix completion

Given a weighted graph, (V, E, δ), find an EDM D such that

$$D_{ij} \approx \delta_{ij}, \quad \forall (i, j) \in E$$

Euclidean distance matrix completion

Undirected graph: $G = (V, E)$ where $V = \{1, \ldots, n\}$.

δ_{ij} = dissimilarity between i and j, available when $(i, j) \in E$.

$D = (d_{ij}) \in \mathbb{R}^{n \times n}$ is a Euclidean distance matrix (EDM) if there is $y_1, \ldots, y_n$ in some Euclidean space such that

$$d_{ij} = \|y_i - y_j\|, \quad \forall i, j \in V$$

Euclidean distance matrix completion

> Given a weighted graph, (V, E, δ), find an EDM D such that
>
> $$D_{ij} \approx \delta_{ij}, \quad \forall (i, j) \in E$$

For example, we could (try to) solve

$$\min_{D \in \mathbb{EDM}} \sum_{(i,j) \in E} (D_{ij} - \delta_{ij})^2$$

Graph embedding

EDM completion is essentially equivalent to the problem of embedding a graph in a Euclidean space...

Graph embedding

EDM completion is essentially equivalent to the problem of embedding a graph in a Euclidean space...

Graph Embedding

Given a weighted graph, (V, E, δ), and embedding dimension d, find $y_1, \ldots, y_n \in \mathbb{R}^d$ such that

$$\|y_i - y_j\| \approx \delta_{ij}, \quad \forall (i,j) \in E$$

Graph embedding

EDM completion is essentially equivalent to the problem of embedding a graph in a Euclidean space...

Graph Embedding

Given a weighted graph, (V, E, δ), and embedding dimension d, find $y_1, \ldots, y_n \in \mathbb{R}^d$ such that

$$\|y_i - y_j\| \approx \delta_{ij}, \quad \forall (i,j) \in E$$

For example, we could (try to) solve

$$\min_{y_1,\ldots,y_n \in \mathbb{R}^d} \sum_{(i,j)\in E} \left(\|y_i - y_j\| - \delta_{ij} \right)^2$$

(Known as non-metric scaling in the statistics literature.)

The problem is also known as multidimensional scaling, graph drawing, graph realization, sensor localization, etc

The problem is also known as multidimensional scaling, graph drawing, graph realization, sensor localization, etc

Important connections to:

- nearest neighbor search
- embedding a finite metric space into a given Banach space

When the graph is complete and there is an exact solution, Classical Scaling finds that solution (by solving an eigenvalue problem).

When the graph is complete and there is an exact solution, Classical Scaling finds that solution (by solving an eigenvalue problem).

It is known to be robust to noise.
(Arias-Castro, Javanmard, and Pelletier 2018)

In the general case, some dissimilarities are missing...

[1] Thorpe and Duxbury 1999; Asimow and Roth 1978; Laman 1970.

In the general case, some dissimilarities are missing...

Whether a graph can be uniquely embedded is a central question in rigidity theory.[1]

[1] Thorpe and Duxbury 1999; Asimow and Roth 1978; Laman 1970.

A number of methods have been proposed:

[2] Kruskal and Seery 1980; Shang et al. 2003; Shang and Ruml 2004; Niculescu and Nath 2003.

[3] Eren et al. 2004; Krislock and Wolkowicz 2010.

[4] Koren, Gotsman, and Ben-Chen 2005; Cucuringu, Lipman, and Singer 2012; Singer 2008; Zhang et al. 2010; Drusvyatskiy et al. 2017.

[5] Leeuw and Mair 2009.

[6] Biswas et al. 2006; Javanmard and Montanari 2013; Weinberger et al. 2006; So and Ye 2007.

A number of methods have been proposed:

- Use the graph distances to fill in the missing dissimilarities and apply classical scaling.[2]

[2] Kruskal and Seery 1980; Shang et al. 2003; Shang and Ruml 2004; Niculescu and Nath 2003.

[3] Eren et al. 2004; Krislock and Wolkowicz 2010.

[4] Koren, Gotsman, and Ben-Chen 2005; Cucuringu, Lipman, and Singer 2012; Singer 2008; Zhang et al. 2010; Drusvyatskiy et al. 2017.

[5] Leeuw and Mair 2009.

[6] Biswas et al. 2006; Javanmard and Montanari 2013; Weinberger et al. 2006; So and Ye 2007.

A number of methods have been proposed:

- Use the graph distances to fill in the missing dissimilarities and apply classical scaling.[2]

- Embed a clique by classical scaling and then sequentially position the nodes for which this is possible.[3]

[2] Kruskal and Seery 1980; Shang et al. 2003; Shang and Ruml 2004; Niculescu and Nath 2003.

[3] Eren et al. 2004; Krislock and Wolkowicz 2010.

[4] Koren, Gotsman, and Ben-Chen 2005; Cucuringu, Lipman, and Singer 2012; Singer 2008; Zhang et al. 2010; Drusvyatskiy et al. 2017.

[5] Leeuw and Mair 2009.

[6] Biswas et al. 2006; Javanmard and Montanari 2013; Weinberger et al. 2006; So and Ye 2007.

A number of methods have been proposed:

- Use the graph distances to fill in the missing dissimilarities and apply classical scaling.[2]
- Embed a clique by classical scaling and then sequentially position the nodes for which this is possible.[3]
- Embed all cliques using classical scaling and synchronize the resulting embedded point sets.[4]

---

[2] Kruskal and Seery 1980; Shang et al. 2003; Shang and Ruml 2004; Niculescu and Nath 2003.

[3] Eren et al. 2004; Krislock and Wolkowicz 2010.

[4] Koren, Gotsman, and Ben-Chen 2005; Cucuringu, Lipman, and Singer 2012; Singer 2008; Zhang et al. 2010; Drusvyatskiy et al. 2017.

[5] Leeuw and Mair 2009.

[6] Biswas et al. 2006; Javanmard and Montanari 2013; Weinberger et al. 2006; So and Ye 2007.

A number of methods have been proposed:

- Use the graph distances to fill in the missing dissimilarities and apply classical scaling.[2]
- Embed a clique by classical scaling and then sequentially position the nodes for which this is possible.[3]
- Embed all cliques using classical scaling and synchronize the resulting embedded point sets.[4]
- Solve by direct optimization via majorization.[5]

---

[2] Kruskal and Seery 1980; Shang et al. 2003; Shang and Ruml 2004; Niculescu and Nath 2003.

[3] Eren et al. 2004; Krislock and Wolkowicz 2010.

[4] Koren, Gotsman, and Ben-Chen 2005; Cucuringu, Lipman, and Singer 2012; Singer 2008; Zhang et al. 2010; Drusvyatskiy et al. 2017.

[5] Leeuw and Mair 2009.

[6] Biswas et al. 2006; Javanmard and Montanari 2013; Weinberger et al. 2006; So and Ye 2007.

A number of methods have been proposed:

- Use the graph distances to fill in the missing dissimilarities and apply classical scaling.[2]
- Embed a clique by classical scaling and then sequentially position the nodes for which this is possible.[3]
- Embed all cliques using classical scaling and synchronize the resulting embedded point sets.[4]
- Solve by direct optimization via majorization.[5]
- Solve a semidefinite program after an appropriate relaxation.[6]

---

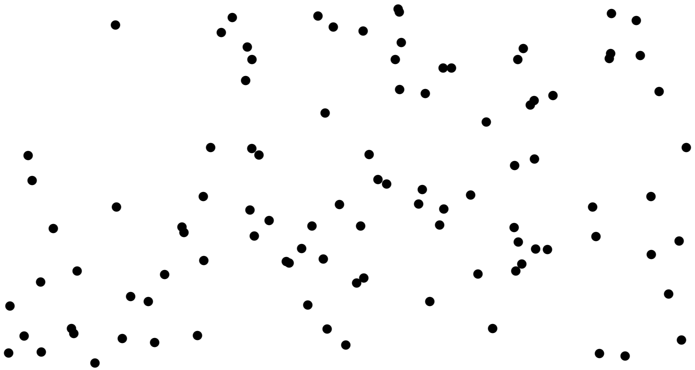
[2] Kruskal and Seery 1980; Shang et al. 2003; Shang and Ruml 2004; Niculescu and Nath 2003.

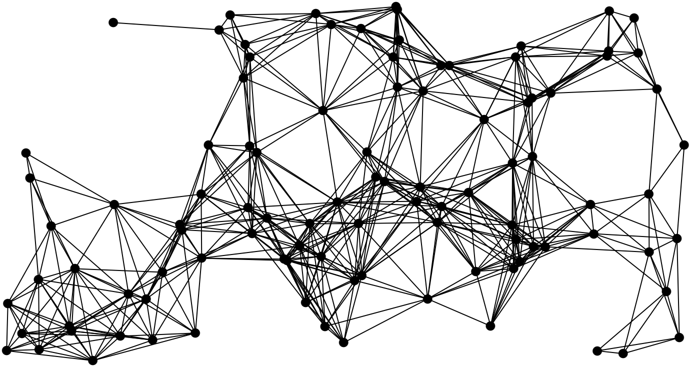
[3] Eren et al. 2004; Krislock and Wolkowicz 2010.

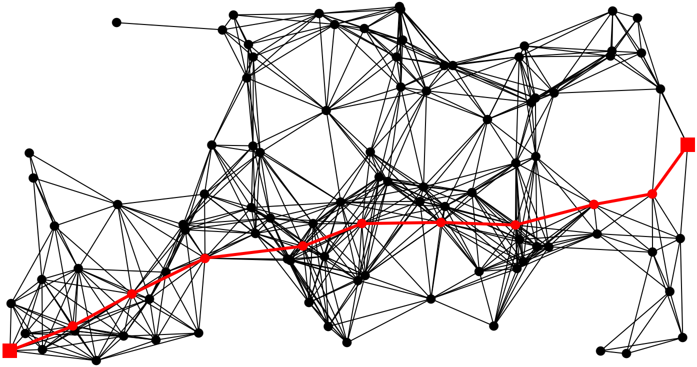
[4] Koren, Gotsman, and Ben-Chen 2005; Cucuringu, Lipman, and Singer 2012; Singer 2008; Zhang et al. 2010; Drusvyatskiy et al. 2017.

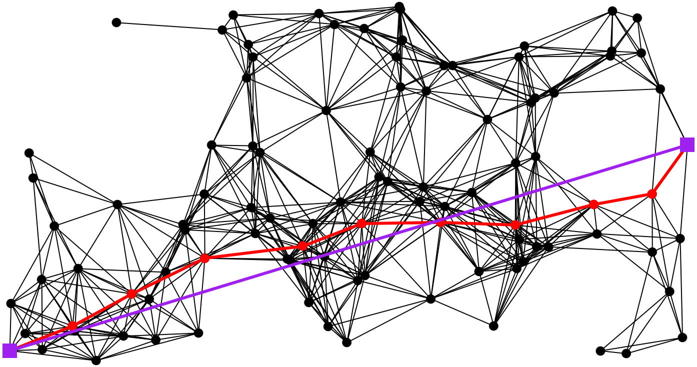
[5] Leeuw and Mair 2009.

[6] Biswas et al. 2006; Javanmard and Montanari 2013; Weinberger et al. 2006; So and Ye 2007.









Graph distances

Based on δ, define the graph distances

$$\Delta(i,j) = \inf_{k_1,\dots,k_m} \sum_{s=0}^{m} \delta(k_s, k_{s+1}),$$

where the infimum is over paths $(k_0, \dots, k_{m+1})$ with $k_0 = i$ and $k_{m+1} = j$.

Bound for neighborhood graphs

Suppose that the graph is in fact the r-ball neighborhood graph of a set of points $x_1, \dots, x_n \in \mathbb{R}^d$, meaning

$$(i, j) \in E \iff \delta_{ij} = \|x_i - x_j\| \le r$$

[7] Arias-Castro, Javanmard, and Pelletier 2018.

Bound for neighborhood graphs

Suppose that the graph is in fact the r-ball neighborhood graph of a set of points $x_1, \dots, x_n \in \mathbb{R}^d$, meaning

$$(i,j) \in E \;\Leftrightarrow\; \delta_{ij} = \|x_i - x_j\| \leq r$$

Proposition[7]

Take $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$ and define

$$\varepsilon = \max_{x \in \mathrm{Conv}(\mathcal{X})} \min_{i \in [n]} \|x - x_i\|$$

When $\varepsilon / r \leq 1/c_1$,

$$\|x_i - x_j\| \leq \Delta(i,j) \leq (1 + c_2(\varepsilon/r)^2)\|x_i - x_j\|$$

where c_1, c_2 are universal constants.

[7] Arias-Castro, Javanmard, and Pelletier 2018.

Bernstein et al. 2000 prove a weaker bound... but in the context of manifold learning, when analyzing Isomap in the setting of a convex domain isometrically embedded in higher dimensions.

Bernstein et al. 2000 prove a weaker bound... but in the context of manifold learning, when analyzing Isomap in the setting of a convex domain isometrically embedded in higher dimensions.

Oh, Montanari, and Karbasi 2010 prove a similar bound in the context of graph drawing for essentially the same method (MDS-MAP of Shang et al. 2003).

Bernstein et al. 2000 prove a weaker bound... but in the context of manifold learning, when analyzing Isomap in the setting of a convex domain isometrically embedded in higher dimensions.

Oh, Montanari, and Karbasi 2010 prove a similar bound in the context of graph drawing for essentially the same method (MDS-MAP of Shang et al. 2003).

(The same bound above holds in that context too.)

Estimating the shortest paths distances on a surface

In manifold learning, the distances of interest are the intrinsic distances on the underlying surface.[8]

[8] Silva and Tenenbaum 2002; Tenenbaum, Silva, and Langford 2000.

[9] Karaman and Frazzoli 2011; Karaman and Frazzoli 2010; Janson et al. 2015; Schmerling, Janson, and Pavone 2015a; Schmerling, Janson, and Pavone 2015b.

In manifold learning, the distances of interest are the intrinsic distances on the underlying surface.[8]

The same is true in motion planning.[9]

[8] Silva and Tenenbaum 2002; Tenenbaum, Silva, and Langford 2000.

[9] Karaman and Frazzoli 2011; Karaman and Frazzoli 2010; Janson et al. 2015; Schmerling, Janson, and Pavone 2015a; Schmerling, Janson, and Pavone 2015b.

Shortest paths

Consider a subset $\mathcal{S} \subset \mathbb{R}^D$.

Shortest paths

Consider a subset $\mathcal{S} \subset \mathbb{R}^D$.

The intrinsic distance on $\mathcal{S}$ is defined, for $x, x' \in \mathcal{S}$, as

$$g(x, x') = \inf \Big\{ a : \exists \gamma : [0, a] \to \mathcal{S},\ 1\text{-Lipschitz}, \\ \text{with } \gamma(0) = x \text{ and } \gamma(a) = x' \Big\}$$

We have a sample of points $\mathcal{X} = \{x_1, \ldots, x_n\} \subset \mathcal{S}$, where $\mathcal{S}$ is compact and connected in $\mathbb{R}^D$. The goal is to estimate $g(x_i, x_j)$.

We have a sample of points $\mathcal{X} = \{x_1, \ldots, x_n\} \subset \mathcal{S}$, where $\mathcal{S}$ is compact and connected in $\mathbb{R}^D$. The goal is to estimate $g(x_i, x_j)$.

Let $\Delta(x, x')$ denote the distance of $x, x' \in \mathcal{X}$ in the r-ball neighborhood graph built on $\mathcal{X}$.

We have a sample of points $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathcal{S}$, where $\mathcal{S}$ is compact and connected in $\mathbb{R}^D$. The goal is to estimate $g(x_i, x_j)$.

Let $\Delta(x, x')$ denote the distance of $x, x' \in \mathcal{X}$ in the r-ball neighborhood graph built on $\mathcal{X}$.

Define

$$\varepsilon = \sup_{x \in \mathcal{S}} \min_{i \in [n]} \|x - x_i\|.$$

We have a sample of points $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathcal{S}$, where $\mathcal{S}$ is compact and connected in $\mathbb{R}^D$. The goal is to estimate $g(x_i, x_j)$.

Let $\Delta(x, x')$ denote the distance of $x, x' \in \mathcal{X}$ in the r-ball neighborhood graph built on $\mathcal{X}$.

Define

$$\varepsilon = \sup_{x \in \mathcal{S}} \min_{i \in [n]} \|x - x_i\|.$$

Proposition (Bernstein et al. 2000)

When $\varepsilon \leq r/4$, we have

$$\Delta(x, x') \leq (1 + 4\varepsilon/r) g(x, x'), \quad \forall x, x' \in \mathcal{X}.$$

Assume that

- The intrinsic and ambient topologies coincide on $\mathcal{S}$.
- The shortest paths on $\mathcal{S}$ have curvature bounded by κ.

Assume that

- The intrinsic and ambient topologies coincide on $\mathcal{S}$.
- The shortest paths on $\mathcal{S}$ have curvature bounded by κ.

Proposition (Bernstein et al. 2000; Arias-Castro and Le Gouic 2017)

There is τ depending on (the reach of) $\mathcal{S}$ and c_0 universal such that, when $r \leq \tau$ and $\kappa r \leq 1/3$,

$$g(x, x') \leq (1 + c_0 r^2)\Delta(x, x'), \quad \forall x, x' \in \mathcal{X}.$$

We also show that every shortest path on $\mathcal{S}$ between two sample points can be approximated by a shortest path in the neighborhood graph...

We also show that every shortest path on $\mathcal{S}$ between two sample points can be approximated by a shortest path in the neighborhood graph... and vice versa.

Curvature-constrained shortest paths

We also study curvature-constrained shortest paths...

Curvature-constrained shortest paths

We also study curvature-constrained shortest paths...

For $\kappa > 0$, the κ-curvature-constrained intrinsic semi-distance on $\mathcal{S}$ is defined, for $x, x' \in \mathcal{S}$, as

$$g_\kappa(x, x') = \inf \Big\{ a : \text{there is } \gamma \text{ as before} \\ \text{with curvature bounded by } \kappa \Big\}$$

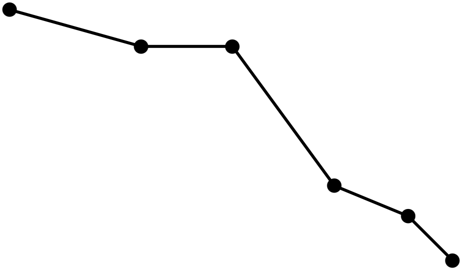
We need a notion of curvature for polygonal lines (which is how paths in a neighborhood graph are embedded).

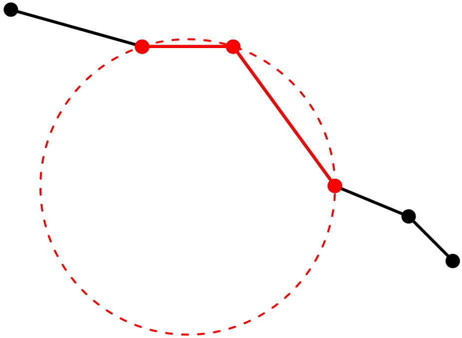
We need a notion of curvature for polygonal lines (which is how paths in a neighborhood graph are embedded).

For an ordered triplet of points (x, y, z) in $\mathbb{R}^D$, define its angle as $\angle(x, y, z) = \angle(\vec{yx}, \vec{yz}) \in [0, \pi]$ and its curvature as

$$\operatorname{curv}(x, y, z) = \begin{cases} 1/R(x, y, z), & \text{if } \angle(x, y, z) \geq \frac{\pi}{2}, \\ \infty, & \text{otherwise,} \end{cases}$$

where $R(x, y, z)$ is the radius of the circle passing through x, y, z.





There are other notions of discrete curvature[10]. This one is consistent in the following sense.

[10] Hoffmann 2009.

There are other notions of discrete curvature[10]. This one is consistent in the following sense.

Lemma

Consider a curve $\gamma : (a, b) \to \mathbb{R}^D$ which is twice continuously differentiable. Holding $s \in (a, b)$ fixed while $r \nearrow s$ and $t \searrow s$,

$$\mathrm{curv}(\gamma(r), \gamma(s), \gamma(t)) \to \text{curvature of } \gamma \text{ at } s.$$

[10] Hoffmann 2009.

We also have the following key lemma.

We also have the following key lemma.

Lemma

Let γ be a simple curve with curvature at most κ. If $x, y, z \in \gamma$ are such that y is between x and z on γ and $\|x - z\| \leq 2/\kappa$, then

$$\mathrm{curv}(x, y, z) \leq \kappa$$

For technical reasons, we work with an annulus graph, where edges with distance $\leq r/4$ are removed from the r-ball graph.

For technical reasons, we work with an annulus graph, where edges with distance $\leq r/4$ are removed from the r-ball graph.

Let $\Delta_\kappa(x, x')$ now denote the length of the shortest path in the graph with curvature bounded by κ.

Proposition

There is a numerical constant $c \geq 1$ such that, when $\varepsilon/r \leq 1/c$, $\kappa r \leq 1/c$, and $\kappa' \geq \kappa + c(\kappa^2 r + \varepsilon/r^2)$,

$$\Delta_\kappa(x, x') \leq (1 + 6\varepsilon/r) g_\kappa(x, x'), \quad x, x' \in \mathcal{X}$$

(The right-hand side may be infinite.)

We now assume in addition that all the shortest paths on $\mathcal{S}$ have curvature bounded by κ.

[11] Alexander and Alexander 1981; Alexander, Berg, and Bishop 1987; Albrecht and Berg 1991.

We now assume in addition that all the shortest paths on $\mathcal{S}$ have curvature bounded by κ.

Lemma

Assume that $\mathcal{S}$ is a compact and connected C^2 submanifold with boundary that is either empty or C^2. Then there is $\kappa < \infty$ such that all the shortest paths on $\mathcal{S}$ have max-curvature bounded by κ.

(Strange things near the boundary.[11])

[11] Alexander and Alexander 1981; Alexander, Berg, and Bishop 1987; Albrecht and Berg 1991.

Theorem

There is a universal constant $c > 0$ such that, if $\kappa r \leq 1/c$ and $\varepsilon/\kappa r^2 \leq 1/c$, the unconstrained shortest paths in the graph have curvature at most $\kappa' \leq \kappa + c\,\varepsilon/\kappa r^3$.

Estimating distances based on adjacency information

Latent graph models

We observe the adjacency matrix $W = (W_{ij})$ of an undirected graph. We assume the existence of points, $x_1, \ldots, x_n \in \mathbb{R}^v$, such that

$$\mathbb{P}(W_{ij} = 1 \mid x_1, \ldots, x_n) = \phi(\|x_i - x_j\|),$$

for some non-increasing link function $\phi : [0, \infty) \mapsto [0, 1]$.

Latent graph models

We observe the adjacency matrix $W = (W_{ij})$ of an undirected graph. We assume the existence of points, $x_1, \dots, x_n \in \mathbb{R}^v$, such that

$$\mathbb{P}(W_{ij} = 1 \mid x_1, \dots, x_n) = \phi(\|x_i - x_j\|),$$

for some non-increasing link function $\phi : [0, \infty) \mapsto [0, 1]$.

- The $(W_{ij}, i < j)$ are assumed to be independent.

Latent graph models

We observe the adjacency matrix $W = (W_{ij})$ of an undirected graph. We assume the existence of points, $x_1, \dots, x_n \in \mathbb{R}^v$, such that

$$\mathbb{P}(W_{ij} = 1 \mid x_1, \dots, x_n) = \phi(\|x_i - x_j\|),$$

for some non-increasing link function $\phi : [0, \infty) \mapsto [0, 1]$.

- The $(W_{ij}, i < j)$ are assumed to be independent.
- The link function ϕ may be known or unknown.

Latent graph models

We observe the adjacency matrix $W = (W_{ij})$ of an undirected graph. We assume the existence of points, $x_1, \ldots, x_n \in \mathbb{R}^v$, such that

$$\mathbb{P}(W_{ij} = 1 \mid x_1, \ldots, x_n) = \phi(\|x_i - x_j\|),$$

for some non-increasing link function $\phi : [0, \infty) \mapsto [0, 1]$.

- The $(W_{ij}, i < j)$ are assumed to be independent.
- The link function ϕ may be known or unknown.

Our goal is to estimate the pairwise distances

$$d_{ij} := \|x_i - x_j\|.$$

Hoff, Raftery, and Handcock 2002 assume a parametric model.

[12] Sarkar, Chakrabarti, and Moore 2010; Liben-Nowell and Kleinberg 2007; Liben-Nowell and Kleinberg 2003.

[13] Sussman, Tang, and Priebe 2014; Tang, Sussman, and Priebe 2013.

[14] Alamgir and Luxburg 2012; Luxburg and Alamgir 2013.

Hoff, Raftery, and Handcock 2002 assume a parametric model.

Interestingly, there is a close connection with the literature on link prediction[12], where one wants to determine which nodes are closest at a given point in time as they are the most likely to become connected in the near future.

[12] Sarkar, Chakrabarti, and Moore 2010; Liben-Nowell and Kleinberg 2007; Liben-Nowell and Kleinberg 2003.

[13] Sussman, Tang, and Priebe 2014; Tang, Sussman, and Priebe 2013.

[14] Alamgir and Luxburg 2012; Luxburg and Alamgir 2013.

Hoff, Raftery, and Handcock 2002 assume a parametric model.

Interestingly, there is a close connection with the literature on link prediction[12], where one wants to determine which nodes are closest at a given point in time as they are the most likely to become connected in the near future.

There is related work by Carey Priebe et al, some of it in the context of a dot product graph — where $\phi(\|x_i - x_j\|)$ is replaced by $\phi(\langle x_i, x_j \rangle)$.[13]

---

[12] Sarkar, Chakrabarti, and Moore 2010; Liben-Nowell and Kleinberg 2007; Liben-Nowell and Kleinberg 2003.

[13] Sussman, Tang, and Priebe 2014; Tang, Sussman, and Priebe 2013.

[14] Alamgir and Luxburg 2012; Luxburg and Alamgir 2013.

Hoff, Raftery, and Handcock 2002 assume a parametric model.

Interestingly, there is a close connection with the literature on link prediction[12], where one wants to determine which nodes are closest at a given point in time as they are the most likely to become connected in the near future.

There is related work by Carey Priebe et al, some of it in the context of a dot product graph — where $\phi(\|x_i - x_j\|)$ is replaced by $\phi(\langle x_i, x_j \rangle)$.[13]

Ulrike von Luxburg et al have considered the case where, instead, a K-nearest neighbor graph is available.[14]

---

[12] Sarkar, Chakrabarti, and Moore 2010; Liben-Nowell and Kleinberg 2007; Liben-Nowell and Kleinberg 2003.

[13] Sussman, Tang, and Priebe 2014; Tang, Sussman, and Priebe 2013.

[14] Alamgir and Luxburg 2012; Luxburg and Alamgir 2013.

Graph distances

We focus on using graph distances to estimate the underlying Euclidean distances.

Graph distances

We focus on using graph distances to estimate the underlying Euclidean distances.

Let Δ_{ij} denote the distance between i and j in the graph defined by the adjacency matrix W.

Graph distances

We focus on using graph distances to estimate the underlying Euclidean distances.

Let Δ_{ij} denote the distance between i and j in the graph defined by the adjacency matrix W.

First assume that $\phi(d) = \mathbb{I}\{d \leq r\}$ for some $r > 0$. (r may be assumed known without loss of generality.)

Graph distances

We focus on using graph distances to estimate the underlying Euclidean distances.

Let Δ_{ij} denote the distance between i and j in the graph defined by the adjacency matrix W.

First assume that $\phi(d) = \mathbb{I}\{d \leq r\}$ for some $r > 0$. (r may be assumed known without loss of generality.)

We estimate $d_{ij} = \|x_i - x_j\|$ by $\hat{d}_{ij} = r\Delta_{ij}$.

Define

$$\varepsilon = \max_{x \in \mathrm{Conv}(x_1, \ldots, x_n)} \min_{i \in [n]} \|x - x_i\|$$

Theorem

For all $i, j \in [n]$,

$$0 \le \hat{d}_{ij} - d_{ij} \le 4(\varepsilon / r) d_{ij} + r$$

Assume without loss of generality that $r \leq 1/2$.

Assume without loss of generality that $r \leq 1/2$.

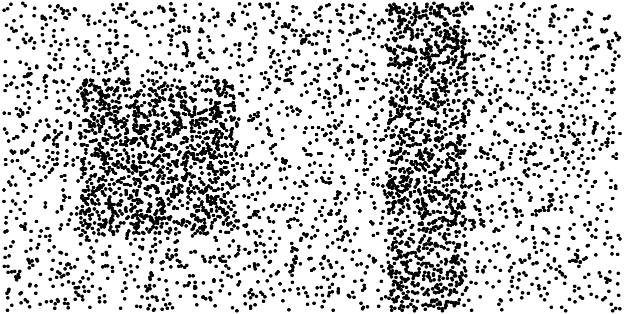
Theorem

There is a numeric constant $c > 0$ with the property that, for any $\varepsilon > 0$ and any estimator $\hat{d}$, there is $x_1, \dots, x_n \in [0,1]$ such that

$$\max_{x \in [0,1]} \min_{i \in [n]} \|x - x_i\| \leq \varepsilon$$

and, for at least half of the pairs $i \neq j$,

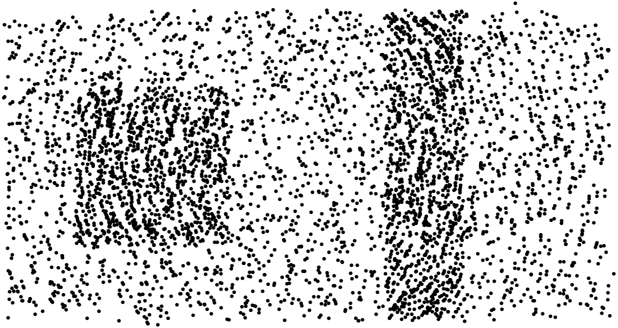
$$|\hat{d}_{ij} - d_{ij}| \geq \frac{c\,\varepsilon}{r \vee \varepsilon} d_{ij}.$$



latent positions (in $[0,2] \times [0,1]$)

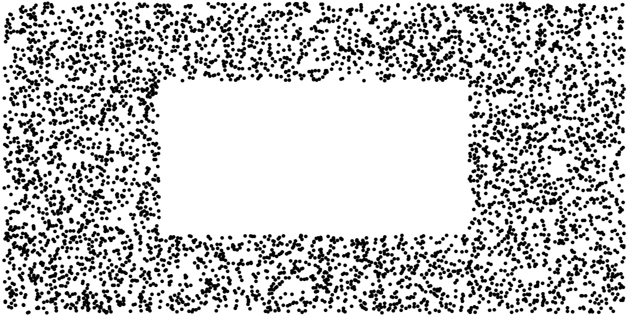
recovered positions with $r = 0.05$

recovered positions with $r = 0.1$

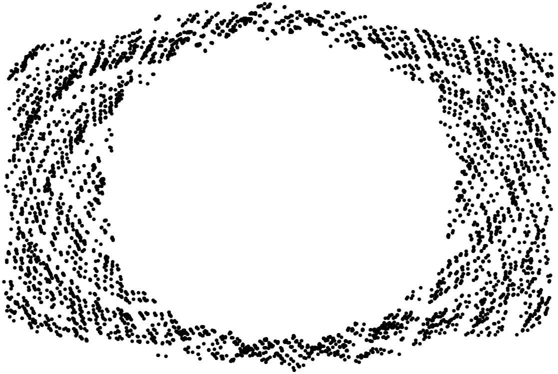


recovered positions with $r = 0.2$

The method requires convexity...



(latent positions)



(estimated positions)

More generally, assume that ϕ has support $[0, r]$, for some $r > 0$, and for some $c_0 > 0$ and $\alpha \geq 0$,

$$\phi(d) \geq c_0(1 - d/r)_+^\alpha$$

(When $\alpha = 0$, ϕ as a discontinuity at $d = r$.)

More generally, assume that ϕ has support $[0, r]$, for some $r > 0$, and for some $c_0 > 0$ and $\alpha \geq 0$,

$$\phi(d) \geq c_0(1 - d/r)_+^\alpha$$

(When $\alpha = 0$, ϕ as a discontinuity at $d = r$.)

Assume without loss of generality that $\mathrm{diam}(x_1, \ldots, x_n) \leq 1$.

Theorem

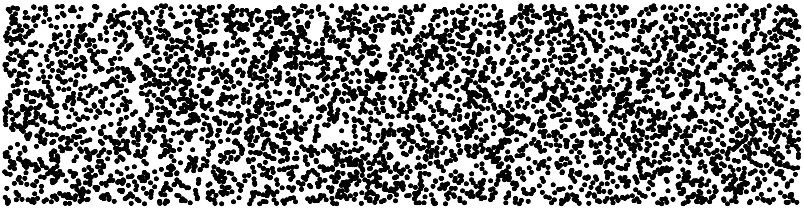
There are $C_1, C_2 > 0$ depending only on (α, c_0) such that, whenever $r/\varepsilon \geq C_1(\log n)^{1+\alpha}$, with probability at least $1-1/n$, for all $i, j \in [n]$,

$$0 \leq \hat{d}_{ij} - d_{ij} \leq C_2\Big[(\varepsilon/r)^{\frac{1}{1+\alpha}} d_{ij} + r\Big]$$

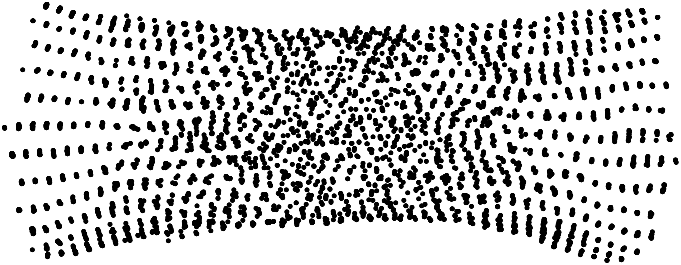
We also obtain results for the setting where the graph is the K-nearest neighbor graph of a point set $x_1, \ldots, x_n$, a setting first considered by Alamgir and Luxburg 2012.

We also obtain results for the setting where the graph is the K-nearest neighbor graph of a point set $x_1, \ldots, x_n$, a setting first considered by Alamgir and Luxburg 2012.

The graph distances perform similarly when $x_1, \ldots, x_n$ are generated iid from the uniform distribution on a compact and convex subset Ω... but only for pairs of points away from the boundary $\partial\Omega$.

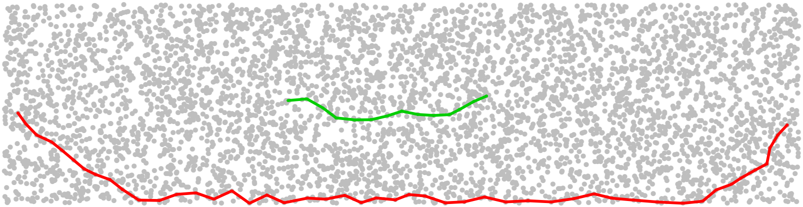


(latent positions)



(estimated positions)

The boundary acts as a high-speed freeway...



Related papers

- Ery Arias-Castro and Bruno Pelletier (2013). "On the convergence of maximum variance unfolding". In: *The Journal of Machine Learning Research* 14.1, pp. 1747–1770
- Ery Arias-Castro and Thibaut Le Gouic (2017). "Unconstrained and Curvature-Constrained Shortest-Path Distances and their Approximation". In: *arXiv preprint arXiv:1706.09441*
- Ery Arias-Castro et al. (2018). "On the Estimation of Latent Distances Using Graph Distances". In: *arXiv preprint arXiv:1804.10611*
- Ery Arias-Castro, Adel Javanmard, and Bruno Pelletier (2018). "Perturbation Bounds for Procrustes, Classical Scaling, and Trilateration, with Applications to Manifold Learning". In: *arXiv preprint arXiv:1810.09569*
- Clément Berenfeld and Ery Arias-Castro (2018). "Some Random Paths with Angle Constraints". In: *arXiv preprint arXiv:1811.01101*

Thank you

Further references I

Alamgir, M and U von Luxburg (2012). "Shortest path distance in random k-nearest neighbor graphs". In: *International Conference on Machine Learning (ICML)*.

Albrecht, Felix and ID Berg (1991). "Geodesics in Euclidean space with analytic obstacle". In: *Proceedings of the American Mathematical Society* 113.1, pp. 201–207.

Alexander, Ralph and Stephanie Alexander (1981). "Geodesics in Riemannian manifolds-with-boundary". In: *Indiana University Mathematics Journal* 30.4, pp. 481–488.

Alexander, Stephanie B, I David Berg, and Richard L Bishop (1987). "The Riemannian obstacle problem". In: *Illinois Journal of Mathematics* 31.1, pp. 167–184.

Arias-Castro, Ery, Adel Javanmard, and Bruno Pelletier (2018). "Perturbation Bounds for Procrustes, Classical Scaling, and Trilateration, with Applications to Manifold Learning". In: *arXiv preprint arXiv:1810.09569*.

Arias-Castro, Ery and Thibaut Le Gouic (2017). "Unconstrained and Curvature-Constrained Shortest-Path Distances and their Approximation". In: *arXiv preprint arXiv:1706.09441*.

Arias-Castro, Ery and Bruno Pelletier (2013). "On the convergence of maximum variance unfolding". In: *The Journal of Machine Learning Research* 14.1, pp. 1747–1770.

Arias-Castro, Ery, Antoine Channarond, Bruno Pelletier, and Nicolas Verzelen (2018). "On the Estimation of Latent Distances Using Graph Distances". In: *arXiv preprint arXiv:1804.10611*.

Asimow, Leonard and Ben Roth (1978). "The rigidity of graphs". In: *Transactions of the American Mathematical Society* 245, pp. 279–289.

Berenfeld, Clément and Ery Arias-Castro (2018). "Some Random Paths with Angle Constraints". In: *arXiv preprint arXiv:1811.01101*.

Bernstein, M., V. De Silva, J.C. Langford, and J.B. Tenenbaum (2000). *Graph approximations to geodesics on embedded manifolds*. Tech. rep. Department of Psychology, Stanford University.

Biswas, Pratik, Tzu-Chen Liang, Kim-Chuan Toh, Yinyu Ye, and Ta-Chung Wang (2006). "Semidefinite programming approaches for sensor network localization with noisy distance measurements". In: *Automation Science and Engineering, IEEE Transactions on* 3.4, pp. 360–371.

Cucuringu, Mihai, Yaron Lipman, and Amit Singer (2012). "Sensor network localization by eigenvector synchronization over the Euclidean group". In: *ACM Transactions on Sensor Networks (TOSN)* 8.3, p. 19.

Drusvyatskiy, Dmitriy, Nathan Krislock, Y-L Voronin, and Henry Wolkowicz (2017). "Noisy Euclidean distance realization: robust facial reduction and the Pareto frontier". In: *SIAM Journal on Optimization* 27.4, pp. 2301–2331.

Further references II

Eren, Tolga, OK Goldenberg, Walter Whiteley, Yang Richard Yang, A Stephen Morse, Brian DO Anderson, and Peter N Belhumeur (2004). "Rigidity, computation, and randomization in network localization". In: *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*. Vol. 4. IEEE, pp. 2673–2684.

Hoff, Peter D, Adrian E Raftery, and Mark S Handcock (2002). "Latent space approaches to social network analysis". In: *Journal of the American Statistical association* 97.460, pp. 1090–1098.

Hoffmann, Tim (2009). *Discrete differential geometry of curves and surfaces*. Vol. 18. Kyushu University.

Janson, Lucas, Edward Schmerling, Ashley Clark, and Marco Pavone (2015). "Fast marching tree: A fast marching sampling-based method for optimal motion planning in many dimensions". In: *The International Journal of Robotics Research* 34.7, pp. 883–921.

Javanmard, Adel and Andrea Montanari (2013). "Localization from incomplete noisy distance measurements". In: *Foundations of Computational Mathematics* 13.3, pp. 297–345.

Karaman, Sertac and Emilio Frazzoli (2010). "Optimal kinodynamic motion planning using incremental sampling-based methods". In: *IEEE Conference on Decision and Control*. IEEE, pp. 7681–7687.

— (2011). "Sampling-based algorithms for optimal motion planning". In: *The International Journal of Robotics Research* 30.7, pp. 846–894.

Koren, Yehuda, Craig Gotsman, and Mirela Ben-Chen (2005). *PATCHWORK: Efficient localization for sensor networks by distributed global optimization*. Tech. rep.

Krislock, Nathan and Henry Wolkowicz (2010). "Explicit sensor network localization using semidefinite representations and facial reductions". In: *SIAM Journal on Optimization* 20.5, pp. 2679–2708.

Kruskal, Joseph B and Judith B Seery (1980). "Designing network diagrams". In: *Conference on Social Graphics*, pp. 22–50.

Laman, Gerard (1970). "On graphs and rigidity of plane skeletal structures". In: *Journal of Engineering mathematics* 4.4, pp. 331–340.

Leeuw, Jan de and Patrick Mair (2009). "Multidimensional Scaling Using Majorization: SMACOF in R". In: *Journal of Statistical Software* 31.i03.

Liben-Nowell, David and Jon Kleinberg (2003). "The Link-Prediction Problem for Social Networks". In: *Conference on Information and Knowledge Management (CIKM'03)*, pp. 556–559.

— (2007). "The link-prediction problem for social networks". In: *Journal of the Association for Information Science and Technology* 58.7, pp. 1019–1031.

Further references III

Luxburg, Ulrike von and Morteza Alamgir (2013). "Density estimation from unweighted k-nearest neighbor graphs: a roadmap". In: *Advances in Neural Information Processing Systems*, pp. 225–233.

Niculescu, Dragoș and Badri Nath (2003). "DV based positioning in ad hoc networks". In: *Telecommunication Systems* 22.1-4, pp. 267–280.

Oh, Sewoong, Andrea Montanari, and Amin Karbasi (2010). "Sensor network localization from local connectivity: Performance analysis for the mds-map algorithm". In: *Information Theory (ITW 2010, Cairo), 2010 IEEE Information Theory Workshop on*. IEEE, pp. 1–5.

Sarkar, Purnamrita, Deepayan Chakrabarti, and Andrew W Moore (2010). "Theoretical Justification of Popular Link Prediction Heuristics". In: *International Conference on Learning Theory*.

Schmerling, Edward, Lucas Janson, and Marco Pavone (2015a). "Optimal sampling-based motion planning under differential constraints: the drift case with linear affine dynamics". In: *IEEE Conference on Decision and Control*. IEEE, pp. 2574–2581.

— (2015b). "Optimal sampling-based motion planning under differential constraints: the driftless case". In: *IEEE International Conference on Robotics and Automation*. IEEE, pp. 2368–2375.

Shang, Yi and Wheeler Ruml (2004). "Improved MDS-based localization". In: *Conference of the IEEE Computer and Communications Societies*. Vol. 4. IEEE, pp. 2640–2651.

Shang, Yi, Wheeler Ruml, Ying Zhang, and Markus PJ Fromherz (2003). "Localization from mere connectivity". In: *ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 201–212.

Silva, V. and J.B. Tenenbaum (2002). "Global versus local methods in nonlinear dimensionality reduction". In: *Advances In Neural Information Processing Systems* 15, pp. 705–712.

Singer, Amit (2008). "A remark on global positioning from local distances". In: *Proceedings of the National Academy of Sciences* 105.28, pp. 9507–9511.

So, Anthony Man-Cho and Yinyu Ye (2007). "Theory of semidefinite programming for sensor network localization". In: *Mathematical Programming* 109.2-3, pp. 367–384.

Sussman, Daniel L, Minh Tang, and Carey E Priebe (2014). "Consistent latent position estimation and vertex classification for random dot product graphs". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.1, pp. 48–57.

Tang, Minh, Daniel L Sussman, and Carey E Priebe (2013). "Universally consistent vertex classification for latent positions graphs". In: *The Annals of Statistics* 41.3, pp. 1406–1430.

Tenenbaum, J. B., V. de Silva, and J. C. Langford (2000). "A Global Geometric Framework for Nonlinear Dimensionality Reduction". In: *Science* 290.5500, pp. 2319–2323.

Further references IV

Thorpe, Michael F and Phillip M Duxbury (1999). *Rigidity theory and applications*. Springer Science & Business Media.

Weinberger, Kilian Q, Fei Sha, Qihui Zhu, and Lawrence K Saul (2006). "Graph Laplacian regularization for large-scale semidefinite programming". In: *Advances in neural information processing systems*, pp. 1489–1496.

Zhang, Lei, Ligang Liu, Craig Gotsman, and Steven J Gortler (2010). "An as-rigid-as-possible approach to sensor network localization". In: *ACM Transactions on Sensor Networks (TOSN)* 6.4, p. 35.