Sensor network localization has benign landscape under mild rank relaxation

June 11, 2024

Chris Criscitiello

with

Andrew McRae, Quentin Rebjock, Nicolas Boumal

OPTIM, Chair of Continuous Optimization Institute of Mathematics, EPFL



$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, ...)

$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, ...)

Will converge to 2-critical point (w/ prob 1)

[1-critical]
$$\nabla f(x) = 0$$

[2-critical]
$$\nabla f(x) = 0$$
 and $\nabla^2 f(x) \ge 0$

$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, ...)

Will converge to 2-critical point (w/ prob 1)

[1-critical] $\nabla f(x) = 0$

[2-critical] $\nabla f(x) = 0$ and $\nabla^2 f(x) \ge 0$

Goal: Show all 2-critical points *x* are global minima.

$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, ...)

Will converge to 2-critical point (w/ prob 1)

[1-critical]
$$\nabla f(x) = 0$$

[2-critical]
$$\nabla f(x) = 0$$
 and $\nabla^2 f(x) \ge 0$

Goal: Show all 2-critical points x are global minima.



$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, ...)

Will converge to 2-critical point (w/ prob 1)

[1-critical] $\nabla f(x) = 0$

[2-critical] $\nabla f(x) = 0$ and $\nabla^2 f(x) \ge 0$

Goal: Show all 2-critical points *x* are global minima.

Benign landscape

The problem

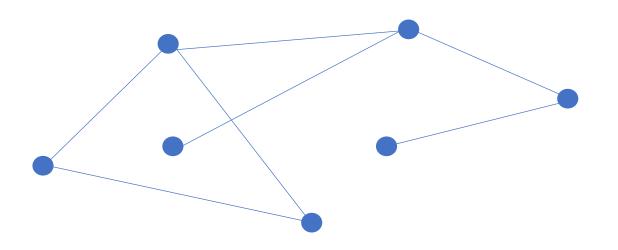
n unknown points $z_1^*, z_2^*, ..., z_n^*$ in \mathbb{R}^{ℓ} .

The problem

n unknown points $z_1^*, z_2^*, ..., z_n^*$ in \mathbb{R}^{ℓ} .

Know a subset of the pairwise distances (measurements)

$$d_{ij} = ||z_i^* - z_j^*|| \text{ for } ij \in E.$$



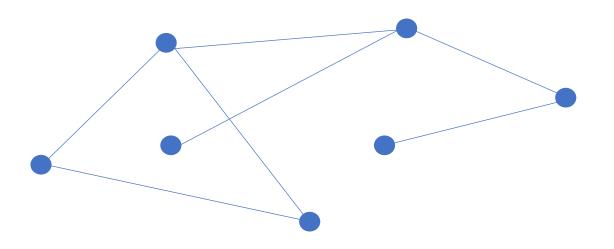
The problem

n unknown points $z_1^*, z_2^*, ..., z_n^*$ in \mathbb{R}^{ℓ} .

Know a subset of the pairwise distances (measurements)

$$d_{ij} = ||z_i^* - z_j^*|| \text{ for } ij \in E.$$

Goal: recover the *n* points (up to translation & rotation)



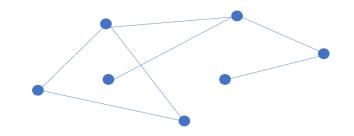
Applications

Robotics (sensor network localization), $\ell = 2,3$

Molecular conformation

Data analysis (metric multidimensional scaling)

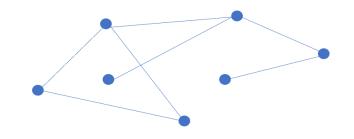
Graph theory (rigidity)



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^{\ell} : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).



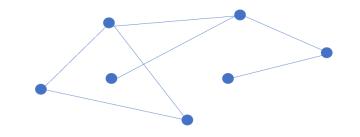
Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^{\ell} : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).

NP-hard!

"Euclidean distance geometry and applications" -- Liberti, et al



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^{\ell} : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).

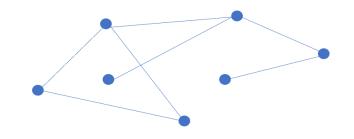
NP-hard!

"Euclidean distance geometry and applications" -- Liberti, et al

Universal rigidity: Configuration space

$$\{z_1, z_2, ..., z_n \in \mathbb{R}^k \text{ for all } k \ge \ell : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^{\ell} : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).

NP-hard!

"Euclidean distance geometry and applications" -- Liberti, et al

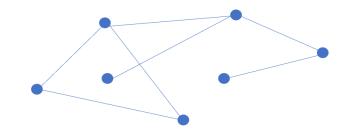
Universal rigidity: Configuration space

$$\{z_1, z_2, ..., z_n \in \mathbb{R}^k \text{ for all } k \ge \ell : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).

Polynomial time by SDPs

"Theory of semidefinite programming for Sensor Network Localization" -- So, Ye



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^{\ell} : d_{ij} = ||z_i - z_j||\}$$

should be a singleton (after quotienting).

NP-hard!

"Euclidean distance geometry and applications" -- Liberti, et al

Universal rigidity: Configuration space

Drawback: SDP involves
$$(n + \ell) \times (n + \ell)$$
 matrices

Polynomial time by SDPs

"Theory of semidefinite programming for Sensor Network Localization" -- So, Ye

Optimization problem

$$\min \sum_{ij \in E} \left(\left\| z_i - z_j \right\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \left\| z_i^* - z_j^* \right\|$$

$$\text{over } z_1, z_2, \dots, z_n \in \mathbb{R}^\ell$$
"Sectrosian"

Optimization problem

$$\min \sum_{ij \in E} \left(\left\| z_i - z_j \right\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \left\| z_i^* - z_j^* \right\|$$
 over $z_1, z_2, \dots, z_n \in \mathbb{R}^\ell$ "s-stress"

Solved via local algorithms. Guarantees?

Nonconvex! How bad?

Optimization problem

$$\min \sum_{ij \in E} \left(\left\| z_i - z_j \right\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \left\| z_i^* - z_j^* \right\|$$
 over $z_1, z_2, \dots, z_n \in \mathbb{R}^\ell$ "s-stress"

Solved via local algorithms. Guarantees?

Nonconvex! How bad?

Possible variations: Noisy measurements, landmarks, ...

Our focus: (nearly) complete graphs, no noise

Synthetic experiments

Recipe (all distances known):

- (1) Choose ground truths $z_1^*, z_2^*, ..., z_n^*$ at random (normal iid)
- (2) Run gradient descent/trust regions/etc.
- (3) Find global min?
- (4) Repeat

Synthetic experiments

Recipe (all distances known):

- (1) Choose ground truths $z_1^*, z_2^*, ..., z_n^*$ at random (normal iid)
- (2) Run gradient descent/trust regions/etc.
- (3) Find global min?
- (4) Repeat

Always finds global min!

Synthetic experiments

Recipe (all distances known):

- (1) Choose ground truths $z_1^*, z_2^*, ..., z_n^*$ at random (normal iid)
- (2) Run gradient descent/trust regions/etc.
- (3) Find global min?
- (4) Repeat

Always finds global min!

Open Question: Does s-stress have spurious local minima? Are all 2-critical points global minima?

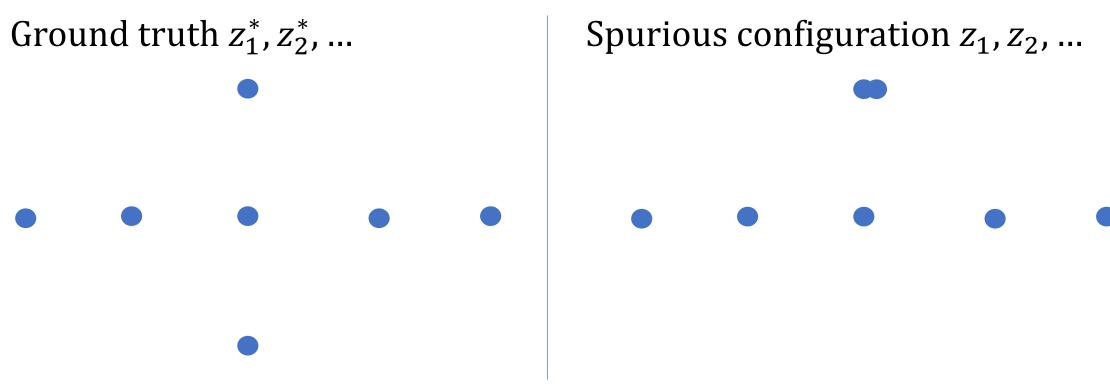
* Malone & Trosset 2000, Parhizkar 2013, etc.

s-stress can have spurious strict local minima!

Ground truth $z_1^*, z_2^*, ...$

 z_2^*, \dots Spurious configuration z_1, z_2, \dots

s-stress can have spurious strict local minima!

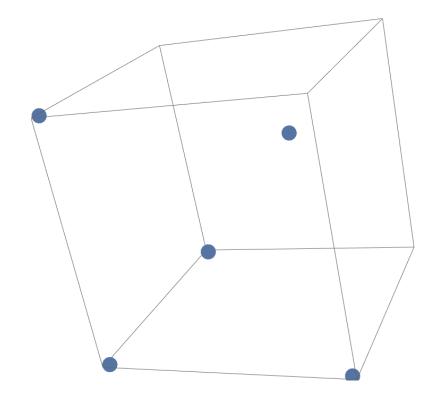


Set of ground truths with spurious local minima has positive measure

Minima number of points to have spurious local minima?

Minima number of points to have spurious local minima?

$$n = \ell + 2 \text{ (for } \ell \geq 5)$$



$$\min \sum_{ij \in E} (\|z_i - z_j\|^2 - d_{ij}^2)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \dots, z_n \in \mathbb{R}^{\ell}$

$$\min \sum_{ij \in E} (\|z_i - z_j\|^2 - d_{ij}^2)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, ..., z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$

$$\min \sum_{ij \in E} (\|z_i - z_j\|^2 - d_{ij}^2)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$
over $z_1, z_2, ..., z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$

Minimizer of relaxed problem same as original? Yes if graph is complete (or more generally if it is *universally rigid*)

$$\min \sum_{ij \in E} \left(\left\| z_i - z_j \right\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \left\| z_i^* - z_j^* \right\|$$

$$\text{over } z_1, z_2, \dots, z_n \in \mathbb{R}^k$$

Relax to dimension $k > \ell$

Minimizer of relaxed problem same as original? Yes if graph is complete (or more generally if it is *universally rigid*)

Want k small; new problem has kn variables

$$\min \sum_{ij \in E} (\|z_i - z_j\|^2 - d_{ij}^2)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$
over $z_1, z_2, ..., z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$

Minimizer of relaxed problem same as original? Yes if graph is complete (or more generally if it is *universally rigid*)

Want k small; new problem has kn variables If k = n - 1, landscape is benign (later)

Can we do better?

Theorem [arbitrary GT]: If graph is complete and relax to $k \approx \ell + \sqrt{n\ell}$,

then every 2-critical point is the ground truth.

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell}$$
,

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic and iid, and relax to

$$k \approx \ell \log(n)$$
,

then every 2-critical point is the ground truth, w.h.p.

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell}$$
,

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic and iid, and relax to

$$k \approx \ell \log(n)$$
,

then every 2-critical point is the ground truth, w.h.p.

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell}$$
,

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic and iid, and relax to

$$k \approx \ell \log(n)$$
,

then every 2-critical point is the ground truth, w.h.p.

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

- True if $n \le \ell + 3$
- Numerical optim to explicitly search for counterexamples.

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell}$$
,

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic and iid, and relax to

$$k \approx \ell \log(n)$$
,

then every 2-critical point is the ground truth, w.h.p.

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

- True if $n \le \ell + 3$
- Numerical optim to explicitly search for counterexamples.

Some ideas from the proof

Ground truth $z_1^*, z_2^*, ...$ in dimension ℓ

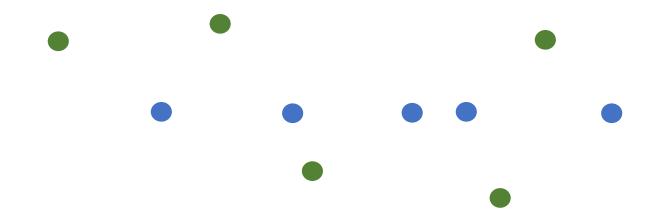
Ground truth $z_1^*, z_2^*, ...$ in dimension ℓ

1-critical configuration in dimension $k > \ell$

Ground truth $z_1^*, z_2^*, ...$ in dimension ℓ

1-critical configuration in dimension $k > \ell$

Goal: perturb 1-critical configuration to decrease cost



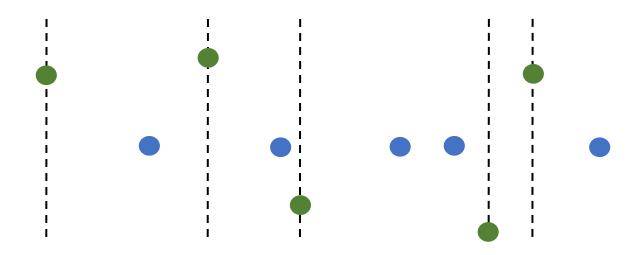
Goal: perturb 1-critical configuration to decrease cost

Goal: perturb 1-critical configuration to decrease cost

$$R = \operatorname{argmin}_{R} \sum_{i} (z_{i}^{*} - Rz_{i})^{2}$$

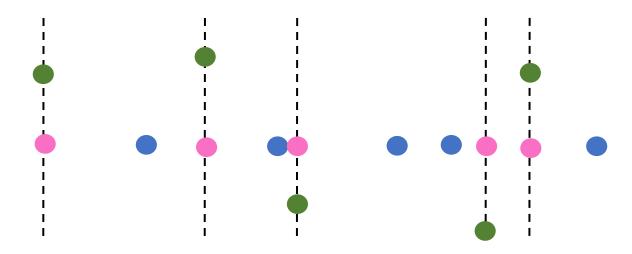
Goal: perturb 1-critical configuration to decrease cost

$$R = \operatorname{argmin}_{R} \sum_{i} (z_{i}^{*} - Rz_{i})^{2}$$



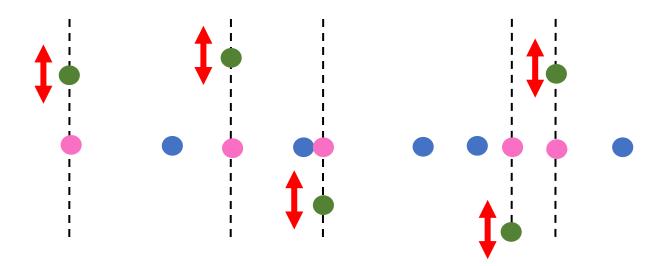
Goal: perturb 1-critical configuration to decrease cost

$$R = \operatorname{argmin}_{R} \sum_{i} (z_{i}^{*} - Rz_{i})^{2}$$

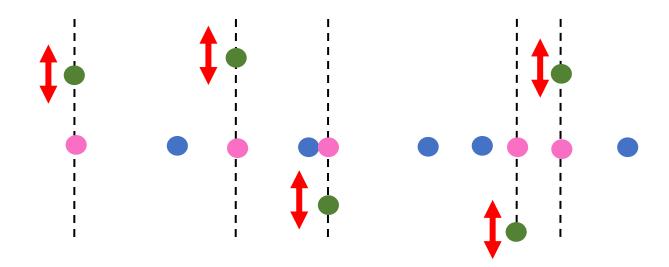


Goal: perturb 1-critical configuration to decrease cost

$$R = \operatorname{argmin}_{R} \sum_{i} (z_{i}^{*} - Rz_{i})^{2}$$

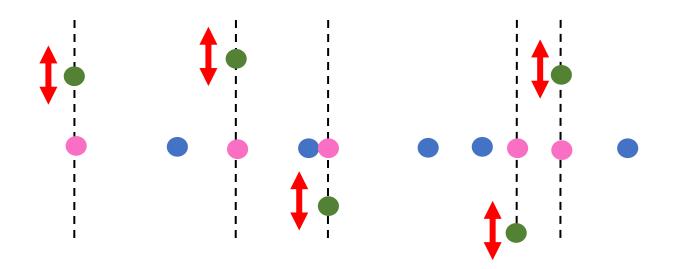


If relax enough, many ways to perturb this way



If relax enough, many ways to perturb this way

Use **eigenvalue interlacing** to argue that a good one exists, if relax enough



If relax enough, many ways to perturb this way

Use **eigenvalue interlacing** to argue that a good one exists, if relax enough

For **isotropic GT**, $k \approx \ell \log(n)$, similar descent direction

Randomize over descent directions (instead of eigenvalue interlacing)

Connection to low-rank optimization, and more

$$Z = \begin{pmatrix} z_1^\mathsf{T} \\ \vdots \\ z_n^\mathsf{T} \end{pmatrix} \in \mathbb{R}^{n \times \ell}, \qquad Z_* = \begin{pmatrix} z_1^{*\mathsf{T}} \\ \vdots \\ z_n^{*\mathsf{T}} \end{pmatrix} \in \mathbb{R}^{n \times \ell}$$

$$Z = \begin{pmatrix} z_1^\mathsf{T} \\ \vdots \\ z_n^\mathsf{T} \end{pmatrix} \in \mathbb{R}^{n \times \ell}, \qquad Z_* = \begin{pmatrix} z_1^{*\mathsf{T}} \\ \vdots \\ z_n^{*\mathsf{T}} \end{pmatrix} \in \mathbb{R}^{n \times \ell}$$

Gram matrices $Y = ZZ^{T}$, $Y_* = Z_*Z_*^{T}$

$$Z = \begin{pmatrix} z_1^\mathsf{T} \\ \vdots \\ z_n^\mathsf{T} \end{pmatrix} \in \mathbb{R}^{n \times \ell}, \qquad Z_* = \begin{pmatrix} z_1^{*\mathsf{T}} \\ \vdots \\ z_n^{*\mathsf{T}} \end{pmatrix} \in \mathbb{R}^{n \times \ell}$$

Gram matrices $Y = ZZ^{T}$, $Y_{*} = Z_{*}Z_{*}^{T}$

MDS map Δ : Sym $(n) \rightarrow \text{Hollow}(n)$

Gram → EDM (euclidean distance matrix)

$$Z = \begin{pmatrix} z_1^\mathsf{T} \\ \vdots \\ z_n^\mathsf{T} \end{pmatrix} \in \mathbb{R}^{n \times \ell}, \qquad Z_* = \begin{pmatrix} z_1^{*\mathsf{T}} \\ \vdots \\ z_n^{*\mathsf{T}} \end{pmatrix} \in \mathbb{R}^{n \times \ell}$$

Gram matrices $Y = ZZ^{\mathsf{T}}$, $Y_* = Z_*Z_*^{\mathsf{T}}$

MDS map Δ : Sym $(n) \rightarrow \text{Hollow}(n)$

Gram → EDM (euclidean distance matrix)

$$[\Delta(Y)]_{ij} := Y_{ii} + Y_{jj} - 2Y_{ij} = ||z_i - z_j||^2$$

 $\min \|\Delta(ZZ^{\mathsf{T}} - Z_*Z_*^{\mathsf{T}})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$

 $\min \|\Delta(ZZ^{\mathsf{T}} - Z_*Z_*^{\mathsf{T}})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$



 $\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geq 0 \text{ with } \operatorname{rank}(Y) \leq \ell$

$$\min \|\Delta(ZZ^{\top} - Z_*Z_*^{\top})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

$$\lim \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geqslant 0 \text{ with } \operatorname{rank}(Y) \leq \ell$$

$$\text{ relax}$$

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geqslant 0 \text{ with } \operatorname{rank}(Y) \leq k$$

$$\min \|\Delta(ZZ^{\top} - Z_*Z_*^{\top})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

$$\downarrow \min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geq 0 \text{ with } \operatorname{rank}(Y) \leq \ell$$

$$\downarrow (\operatorname{relax})$$

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

If k = n, problem is convex

$$\min \|\Delta(ZZ^{\top} - Z_*Z_*^{\top})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

$$\lim \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geq 0 \text{ with } \operatorname{rank}(Y) \leq \ell$$

$$\text{ (relax)}$$

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \geq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

If k = n, problem is convex

Restricted Isometry Property? No! $\Delta^* \circ \Delta$ has condition number n

Special properties of MDS map?

 $\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \ge 0 \text{ with } \operatorname{rank}(Y) \le k$

Special properties of MDS map?

 $\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \ge 0 \text{ with } \operatorname{rank}(Y) \le k$

Special "perturbation" of the identity

$$(\Delta^* \circ \Delta)(Y) = Y + \Theta(Y)$$

$$(\Delta^* \circ \Delta)^{-1}(Y) = Y - \Gamma(Y)$$

Special properties of MDS map?

 $\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \ge 0 \text{ with } \operatorname{rank}(Y) \le k$

Special "perturbation" of the identity

$$(\Delta^* \circ \Delta)(Y) = Y + \Theta(Y)$$
$$(\Delta^* \circ \Delta)^{-1}(Y) = Y - \Gamma(Y)$$

General theorem: If Γ is completely positive, contractive, "behaves well w.r.t. rank-1 matrices", and

$$\langle Y, \Theta(Y) \rangle \leq c \langle Y, \Gamma(Y) \rangle \quad \forall Y$$

then landscape is benign when relax to $k \approx \ell + \sqrt{c\ell}$.

Open questions

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

Incomplete graphs (random, expanders, ...)

Many other localization problems (noise models, trajectory localization, inverse kinemetics, ...)

 $\min \|\Delta(ZZ^{\mathsf{T}} - Z_*Z_*^{\mathsf{T}})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$

min $||ZZ^{\top} - Z_*Z_*^{\top}||^2$ over $Z \in \mathbb{R}^{n \times \ell}$ (with trace $(ZZ^{\top}) = 1$)

$$\min \|\Delta(ZZ^{\mathsf{T}} - Z_*Z_*^{\mathsf{T}})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

```
min ||ZZ^{\top} - Z_*Z_*^{\top}||^2 over Z \in \mathbb{R}^{n \times \ell} (with trace(ZZ^{\top}) = 1)
```

• Kirwan: K = U(n) acts on projective space $\mathbb{P}(\mathbb{C}^{n \times \ell})$

$$\min \|\Delta(ZZ^{\mathsf{T}} - Z_*Z_*^{\mathsf{T}})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

$$\min \|ZZ^{\top} - Z_*Z_*^{\top}\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell} \text{ (with trace}(ZZ^{\top}) = 1)$$

• Kirwan: K = U(n) acts on projective space $\mathbb{P}(\mathbb{C}^{n \times \ell})$

No index-1 critical points if relax to $k = \ell + 2$?

[Index = number of negative eigenvalues of Hessian]

$$\min \|\Delta(ZZ^{\mathsf{T}} - Z_*Z_*^{\mathsf{T}})\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

min
$$||ZZ^{\top} - Z_*Z_*^{\top}||^2$$
 over $Z \in \mathbb{R}^{n \times \ell}$ (with trace $(ZZ^{\top}) = 1$)

• Kirwan: K = U(n) acts on projective space $\mathbb{P}(\mathbb{C}^{n \times \ell})$

No index-1 critical points if relax to $k = \ell + 2$?

Seems to be a common phenomenon when relaxing dimension

[Index = number of negative eigenvalues of Hessian]

Open questions

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

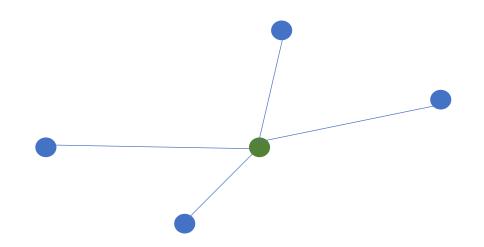
Incomplete graphs (random, expanders, ...)

Many other localization problems (noise models, trajectory localization, inverse kinemetics, ...)

Appendix

SNL with landmarks

$$\min \sum_i \left(\|z - z_i\|^2 - d_i^2 \right)^2, \qquad d_i = \|z^* - z_i^*\|$$
 over $z \in \mathbb{R}^\ell$



SNL with landmarks

$$\min \sum_i \left(\|z - z_i\|^2 - d_i^2 \right)^2, \qquad d_i = \|z^* - z_i^*\|$$
 over $z \in \mathbb{R}^\ell$

Landscape is not benign in general.

SNL with landmarks

$$\min \sum_i \left(\|z-z_i\|^2 - d_i^2 \right)^2, \qquad d_i = \|z^* - z_i^*\|$$
 over $z \in \mathbb{R}^\ell$

Landscape is not benign in general.

Proposition: If relax to $k = \ell + 1$, the landscape is benign.

Hubs

Theorem [isotropic GT]: If graph is **nearly complete**, ground truth points are isotropic and iid, and relax to

$$k \approx \ell \log(n)$$
,

then every 2-critical point is the ground truth.

The **hub** of a graph is the set of vertices which are connected to all other vertices.

$$H = \text{size of hub}$$

Theorem [isotropic GT]: If ground truth points are isotropic and iid, and relax to

$$k \approx \text{poly}(n-H)\ell \log(n)$$
,

then every 2-critical point is the ground truth.