# Lecture 11: Isometric and Nearly Isometric Embeddings of Geometric Graph.

**Radu Balan**

Department of Mathematics, AMSC, CSCAMM and NWC
University of Maryland, College Park, MD

May 2, 2019

# Isometric Embeddings with Full Data

## Problem statement and Ambiguities

### Main Problem

*Isometric Embedding: Given the set of all squared-distances* $\{d_{i,j}^2;\ 1 \leq i,j \leq n\}$ *find a dimension* $d$ *and a set of* $n$ *points* $\{y_1, \cdots, y_n\} \subset \mathbb{R}^d$ *so that* $\|y_i - y_j\|^2 = d_{i,j}^2$, $1 \leq i,j \leq n$ *.*

### Main Problem

*Nearly Isometric Embedding: Given the set of all squared-distances* $\{d_{i,j}^2;\ 1 \leq i,j \leq n\}$ *find a dimension* $d$ *and a set of* $n$ *points* $\{y_1, \cdots, y_n\} \subset \mathbb{R}^d$ *so that* $\|y_i - y_j\|^2 \approx d_{i,j}^2$, $1 \leq i,j \leq n$ *.*

Note the set of points is unique up to rigid transformations: translations, rotations and reflections: $\mathbb{R}^d \times O(d)$. This means two sets of $n$ points in $\mathbb{R}^d$ have the same pairwise distances if and only if one set is obtained from the other set by a combination of rigid transformations.

# Isometric Embeddings with Full Data

## Converting pairwise distances into the Gram matrix

Let $S = (S_{i,j})_{1 \leq i,j \leq n}$ denote the $n \times n$ symmetric matrix of squared pairwise distances:

$$S_{i,j} = d_{i,j}^2 \quad , S_{i,i} = 0$$

Denote by $1$ the $n$-vector of 1's (the Matlab $ones(n,1)$). Let $\nu = (\|y_i\|^2)_{1 \leq i \leq n}$ denote the unknown $n$-vector of squared-norms. Finally, let $G = (\langle y_i, y_j \rangle)_{1 \leq i,j \leq n}$ denote the Gram matrix of scalar products between $y_i$ and $y_j$.

We can remove the translation ambiguity by fixing the center:

$$\sum_{i=1}^{n} y_i = 0$$

# Isometric Embeddings with Full Data

## Converting pairwise distances into the Gram matrix

Expand the square:

$d_{i,j}^2 = \|y_i - y_j\|^2 = \|y_i\|^2 + \|y_j\|^2 - 2\langle y_i, y_j\rangle \Rightarrow 2\langle y_i, y_j\rangle = \|y_i\|^2 + \|y_j\|^2 - d_{i,j}^2$

Rewrite the system as:

$$2G = \nu \cdot 1^T + 1 \cdot \nu^T - S \quad (*)$$

The center condition reads: $G \cdot 1 = 0$, which implies:

$$0 = 2n\nu^T \cdot 1 - 1^T \cdot S \cdot 1$$

Let $\rho := \nu^T \cdot 1 = \sum_{i=1}^n \|y_i\|^2$. We obtain:

$$\rho = \frac{1}{2n} 1^T \cdot S \cdot 1 = \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^n d_{i,j}^2$$

$$\nu = \frac{1}{n}(S \cdot 1 - \rho 1) = \frac{1}{n}(S - \rho I) \cdot 1$$

that you substitute back into (*).

# Isometric Embeddings with Full Data

## Converting pairwise squared-distances into the Gram matrix: Algorithm

### Algorithm

*Input: Symmetric matrix of squared pairwise distances* $S = (d_{i,j}^2)_{1 \leq i,j \leq n}$.

1. *Compute:*
$$\rho = \frac{1}{2n} 1^T \cdot S \cdot 1 = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{i,j}^2$$
2. *Set:*
$$\nu = \frac{1}{n}(S \cdot 1 - \rho 1) = \frac{1}{n}(S - \rho I) \cdot 1$$
3. *Compute:*
$$G = \frac{1}{2}\nu \cdot 1^T + \frac{1}{2} 1 \cdot \nu^T - \frac{1}{2} S = \frac{1}{2n}(S - \rho I) 1 \cdot 1^T + \frac{1}{2n} 1 \cdot 1^T (S - \rho I) - \frac{1}{2} S.$$

*Output: Symmetric Gram matrix G*

# Isometric Embeddings with Full Data

## Factorization of the $G$ matrix

In the absence of noise (i.e. if $S_{i,j}$ are indeed the Euclidean distances), the Gram matrix $G$ should have rank $d$, the minimum dimension of the isometric embedding.
If $S$ is noisy, then $G$ has approximate rank $d$.
To find $d$ and $Y$, the matrix of coordinates, perform the eigendecomposition:

$$G = Q\Lambda Q^T$$

where $\Lambda$ is the diagonal matrix of eigenvalues, ordered monotonically decreasing. Choose $d$ as the number of significant positive eigenvalues (i.e. truncate to zero the negative eigenvalues, as well as the smallest positive eigenvalues). Note $G$ has always at least one zero eigenvalue: $rank(G) \leq n-1$.

# Isometric Embeddings with Full Data

## Factorization of the $G$ matrix

Then we obtain an approximate factorization of $G$ (exact in the absence of noise):

$$G \approx Q_1 \Lambda_1 Q_1^T$$

where $Q_1$ is the $n \times d$ submatrix of $Q$ containing the first $d$ columns.
Set $Y = \Lambda_1^{1/2} Q_1^T$, so that $G \approx Y^T Y$.
The $d \times n$ matrix $Y$ contains the embedding vectors $y_1, \cdots, y_n$ as columns:

$$Y = [y_1 | y_2 | \cdots | y_n] .$$

# Isometric Embeddings with Full Data

## Gram matrix factorization: Algorithm

### Algorithm

*Input: Symmetric $n \times n$ Gram matrix $G$.*

1. *Compute the eigendecomposition of $G$, $G = Q\Lambda Q^T$ with diagonal of $\Lambda$ sorted in a descending order;*
2. *Determine the number $d$ of significant positive eigevalues;*
3. *Partition*
$$Q = [Q_1 \quad Q_2] \quad , \text{and} \quad \Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}$$
*where $Q_1$ contains the first $d$ columns of $Q$, and $\Lambda_1$ is the $d \times d$ diagonal matrix of significant positive eigenvalues of $G$.*
4. *Compute:*
$$Y = \Lambda_1^{1/2} Q_1^T$$

*Output: Dimension $d$ and $d \times n$ matrix $Y$ of vectors $Y = [y_1 | \cdots | y_n]$*

# Isometric Embeddings with Partial Data

## Dimension estimation

Consider now the case that only a subset of the pairwise squared-distances are known, indexed by $\Theta$. Assume that only $m$ distances (out of $n(n-1)/2$ possible values) are known – this means the cardinal of $\Theta$ is $m$.

### Remark

*Minimum number of measurements:* $m \geq nd - \frac{d(d+1)}{2}$*, because:* $nd$ *is the number of degrees of freedom (coordinates) needed to describe n points in* $\mathbb{R}^d$*;* $d(d+1)/2$ *is the the dimension of the Lie group of Euclidean transformations: translations in* $\mathbb{R}^d$ *of dimension* $d$ *and orthogonal transformations* $O(d)$ *of dimension* $d(d-1)/2$ *(the dimension of the Lie algebra of anti-symmetric matrices).*

In the absence of noise, for sufficiently large $m$ but less than $n(n-1)/2$, exact (i.e. isometric) embedding is possible.

# Isometric Embeddings with Partial Data

## Linear constraints

Given any set of vectors $\{y_1, \cdots, y_n\}$ and their associated matrix $Y = [y_1 | \cdots | y_n]$ their invariant to the action of the rigid transformations (translations, rotations, and reflections) is the Gram matrix of the centered system:

$$G = (I - \frac{1}{n} 1 \cdot 1^T) Y^T Y (I - \frac{1}{n} 1 \cdot 1^T) =: LY^T YL \quad , \quad L = I - \frac{1}{n} 1 \cdot 1^T.$$

On the other hand, the distance between points $i$ and $j$ can be computed by:

$$d_{i,j}^2 = \|y_i - y_j\|^2 = G_{i,i} - G_{i,j} + G_{j,j} - G_{j,i} = e_{ij}^T G e_{ij}$$

where

$$e_{ij} = \delta_i - \delta_j = [0 \cdots 0\, 1 \cdots -1\, 0 \cdots 0]^T$$

where 1 is on position $i$, $-1$ is on position $j$, and 0 everywhere else.

# Almost Isometric Embeddings with Partial Data

## The SDP Problem

Reference [10] proposes to find the matrix $G$ by solving the following Semi-Definite Program:

$$\begin{array}{cc} \min\limits_{\substack{G = G^T \geq 0 \\ G1 = 0 \\ |\langle G e_{ij}, e_{ij}\rangle - \tilde{d}^2_{i,j}| \leq \varepsilon \ , \ (i,j) \in \Theta}} & trace(G) \end{array}$$

where $\tilde{d}^2_{i,j}$ are noisy estimates $d_{i,j}$ and $\varepsilon$ is the maximum noise level. The trace promotes low rank in this optimization. However, this is basically a feasibility problem: Decrease $\varepsilon$ to the minimum value where a feasible solution exists. With probability 1 that is unique.

How to do this: Use CVX with Matlab.

# Nearly Isometric Embeddings with Partial Data

## Stability to Noise

[10] proves the following stability result in the case of partial measurements. Here we denote $\Theta_r = \{(i,j)\ ,\ \|y_i - y_j\| \leq r\}$ the set of all pairs of points at distance at most $r$.

### Theorem

*Let $\{y_1, \cdots, y_n\}$ be $n$ nodes distributed uniformly at random in the hypercube $[-0.5, 0.5]^d$. Further, assume that we are given noisy measurement of all distances in $\Theta_r$ for some $r \geq 10\sqrt{d}(\log(n)/n)^{1/d}$ and the induced geometric graph of edges is connected. Let $\tilde{d}_{i,j}^2 = d_{i,j}^2 + \nu_{i,j}$ with $|\nu_{i,j}| \leq \varepsilon$. Then with high probability, the error distance between the estimated $\hat{Y} = [\hat{y}_1, | \cdots | \hat{y}_n]$ returned by the SDP-based algorithm and the correct coordinate matrix $Y = [y_1 | \cdots | y_n]$ is upper bounded as*

$$\|L\hat{Y}^T\hat{Y}L - LY^TYL\|_1 \leq C_1 (nr^d)^5 \frac{\varepsilon}{r^4}.$$

## References


- B. Bollobás, **Graph Theory. An Introductory Course**, Springer-Verlag 1979. **99**(25), 15879–15882 (2002).
- S. Boyd, L. Vandenberghe, **Convex Optimization**, available online at: http://stanford.edu/ boyd/cvxbook/
- F. Chung, **Spectral Graph Theory**, AMS 1997.
- F. Chung, L. Lu, The average distances in random graphs with given expected degrees, Proc. Nat.Acad.Sci. 2002.
- F. Chung, L. Lu, V. Vu, The spectra of random graphs with Given Expected Degrees, Internet Math. **1**(3), 257–275 (2004).
- R. Diestel, **Graph Theory**, 3rd Edition, Springer-Verlag 2005.
- P. Erdös, A. Rényi, On The Evolution of Random Graphs

- G. Grimmett, **Probability on Graphs. Random Processes on Graphs and Lattices**, Cambridge Press 2010.
- C. Hoffman, M. Kahle, E. Paquette, Spectral Gap of Random Graphs and Applications to Random Topology, arXiv: 1201.0425 [math.CO] 17 Sept. 2014.
- [10]A. Javanmard, A. Montanari, Localization from Incomplete Noisy Distance Measurements, arXiv:1103.1417, Nov. 2012; also ISIT 2011.
- J. Leskovec, J. Kleinberg, C. Faloutsos, Graph Evolution: Densification and Shrinking Diameters, ACM Trans. on Knowl.Disc.Data, **1**(1) 2007.