The Euclidean Distance Matrix Completion Problem

CJ, W&M

Overview

- (Euclidean) Distance Matrices, theory
- Matrix Completion Problems
- Important Graph Theory
- The Distance Matrix Case

References

a motivation : molecular conformation

Euclidean Distance Matrices

$P_1, P_2, \cdots, P_n \in \mathbb{R}^k$

$\|x\| = (x^T x)^{1/2}, \quad x \in \mathbb{R}^k$

$d(P_i, P_j) = \|P_i - P_j\|$ Note

$D = \left(d(P_i - P_j)^2\right) \in M_n(\mathbb{R})$

$\left.\begin{array}{l}\text{symmetric, nonnegative} \\ \quad \text{0-diagonal ("hollow")}\end{array}\right\}$ Necessary

D is a "distance matrix" iff
such $P_1, P_2, \cdots, P_n$ exist

embedding dimension —
smallest k

- Blumenthal
- Schönberg

Recognizing EDM's

If $D \in M_n(\mathbb{R})$ meets the obvious necessary conditions, TFAE

i) D is an EDM

ii) D is negative (semi-) definite on the orthogonal complement of e = the vector of $\underline{1}$'s.

iii) the Schur complement of

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ in }$$

$$\begin{bmatrix} 0 & e^T \\ e & D \end{bmatrix}$$

is negative definite

In fact, the rank of this Schur complement is the embedding dimension k for the points - if D is an EDM.

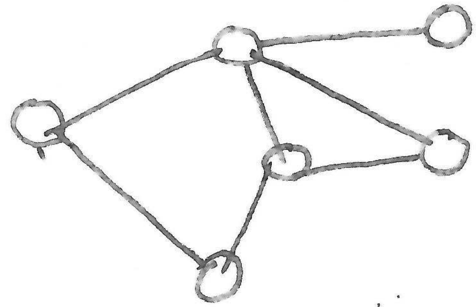
Other Approaches

Linear transformations that that map the EDM's to/from the PD matrices; these transfer the question to checking for PD <u>and</u> produce a set of points, when they exist.

The Important Graph Theory

(Undirected) Graphs:

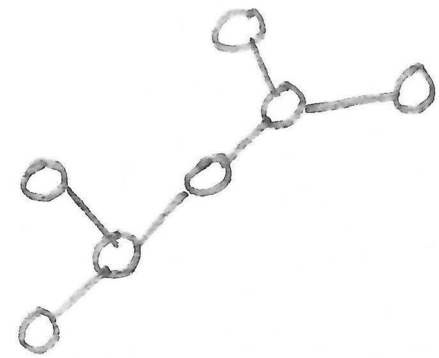
Vertices,
Edges



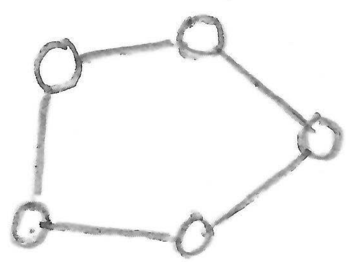
6 V
7 E

Path

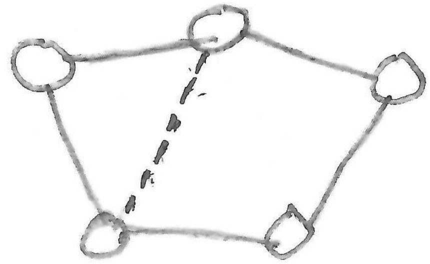
Tree



Cycle

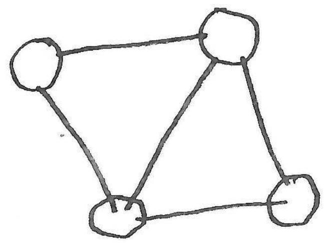
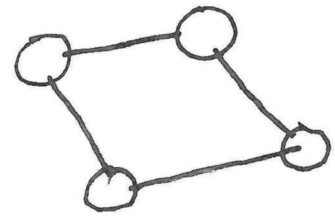


chord
of
a
cycle



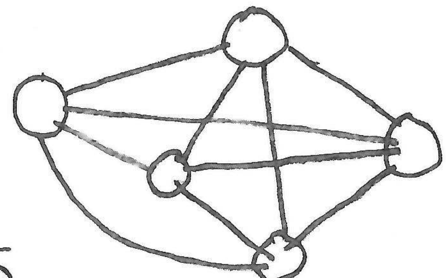
Chordal Graphs connected undirected

- every cycle of length 4 or more has a chord

chordal  Not 

trees are all chordal

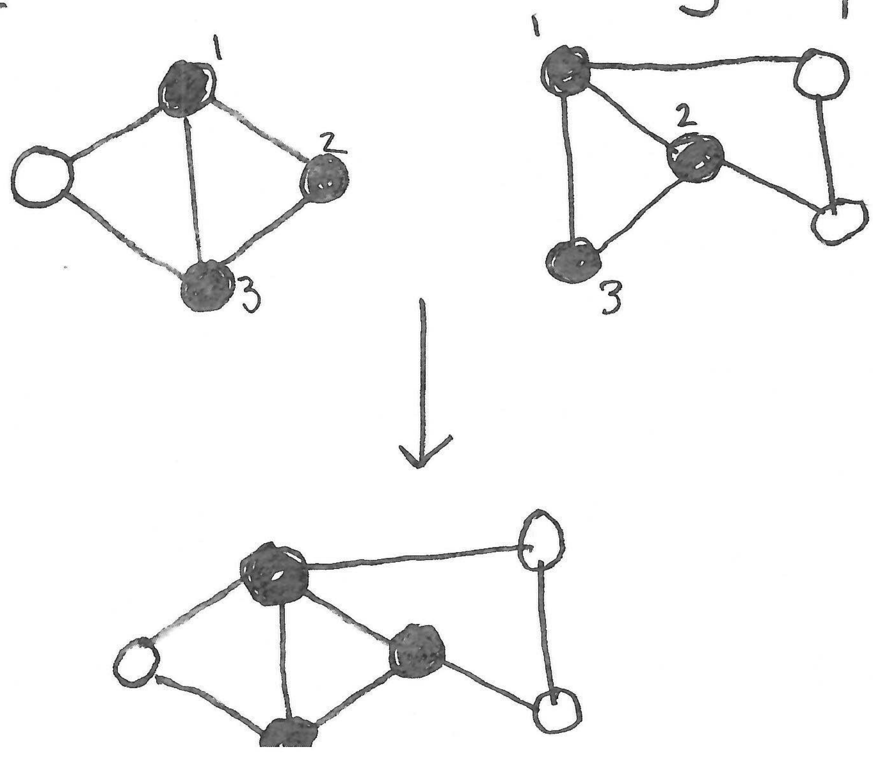
<u>Clique</u>



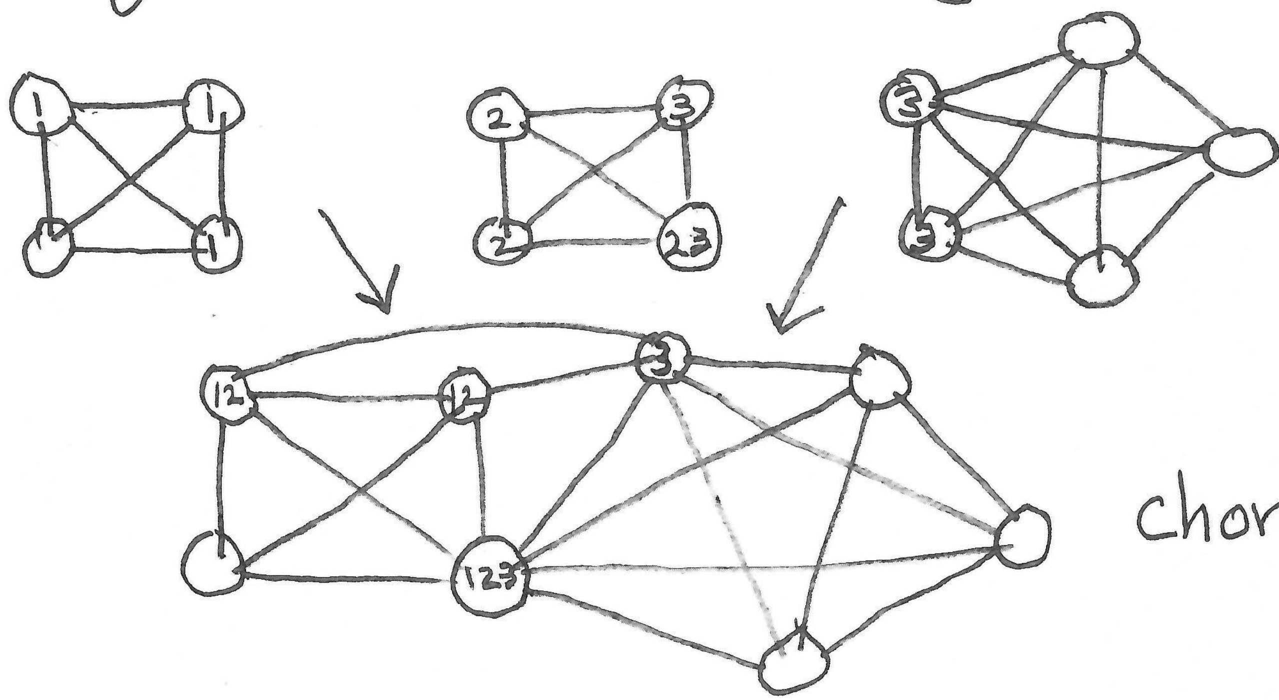
K_5

complete subgraph

clique-sum of 2 graphs



Theorem. A graph is chordal iff it is a sequentia[l] clique-sum of cliques



chordal

Corollary. A graph is chordal iff edges may be added, one at a time, so that the resulting sequence of graphs are all chordal (non-unique)

Matrix Completion Problems

Property $\mathcal{P}$ that a matrix may or may not have

e.g. positive definite (PD)
rank k
Euclidean distance
Stable etc.

For square matrices, $\mathcal{P}$ is said to be inheritted if every principal submatrix of a matrix with $\mathcal{P}$ also has $\mathcal{P}$

e.g. PD, totally positive,
Euclidean distance

Real Polynomially Describable

We say that property $\mathcal{P}$ is RPD if it may be viewed as a semi-algebraic set in $\mathbb{R}^n$ (or $\mathbb{C}^n$), i.e. describable via a finite list of polynomial equalities and inequalities.

Assertion: All properties that anyone cares about are RPD (J, L-D).

in particular, Euclidean distance matrices

"Partial matrix": Some entries <u>specified</u>, while the others are <u>free</u> to be chosen

$$\begin{bmatrix} 3 & -1 & ? & 4 \\ -1 & 5 & 2 & ? \\ ? & 2 & 2 & 0 \\ 4 & ? & 0 & 7 \end{bmatrix}$$

? = free (unspecified)

"Completion": Choice of values for the unspecified entries, resulting in a conventional matrix.

eg $\begin{bmatrix} 3 & -1 & -1 & 4 \\ -1 & 5 & 2 & 1 \\ -1 & 2 & 2 & 0 \\ 4 & 1 & 0 & 7 \end{bmatrix}$

$\mathcal{P}$-completion problem:

Which partial matrices have a completion with property $\mathcal{P}$?

If $\mathcal{P}$ is inheritted, it is necessary that the partial matrix be "partial $\mathcal{P}$", ie. all fully specified principal submatrices have property $\mathcal{P}$.

This necessary condition may or may not be sufficient. It depends on the "pattern" (of specified entries).

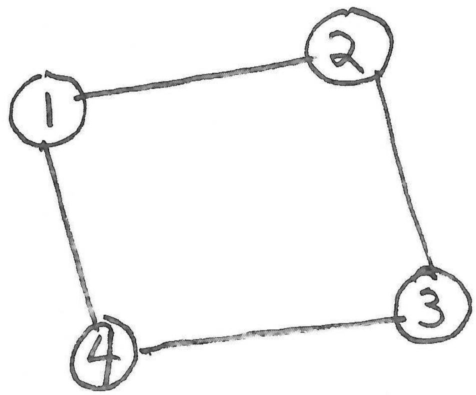
Patterns e.g. $\begin{bmatrix} \bullet & \times & ? & \times \\ \times & \bullet & \times & ? \\ ? & \times & \bullet & \times \\ \times & ? & \times & \bullet \end{bmatrix}$

In the case of combinatorially symmetric patterns, the pattern of the specified entries may be described with a(n) (undirected) graph

i,j entry specified $\longleftrightarrow$

(i) —— (j) is an edge

e.g.



$\begin{bmatrix} \bullet & \times & ? & \times \\ \times & \bullet & \times & ? \\ ? & \times & \bullet & \times \\ \times & ? & \times & \bullet \end{bmatrix}$

Important fact:

If $\mathcal{P}$ is RPD, then, for any fixed pattern, for the specified entries of a partial matrix, the $\mathcal{P}$-completable partial matrices form a semi-algebraic set! So, $\mathcal{P}$-completability may be checked via finitely many polynomial inequalities. (These may not be easy to find and there may be a lot of them.)

For some patterns, and certain inheritted properties $\mathcal{P}$, these conditions are no more than being

partial $\mathcal{P}$!

These $\mathcal{P}$-completable patterns are important to understand

<u>Note.</u> There is a similar discussion for properties involving non-combinatorially-symmetric matrices, or even non-square matrices.

E.g. totally positive (TP) matrices

Classical Result (GJSW 1984)

Every partial positive definite matrix, the graph of whose specified entries is G, has a positive definite completion iff G is chordal!

much more to say here

The EDM Completion Problem:

Which partial matrices have an EDM completion ?

EDM : an inheritted property

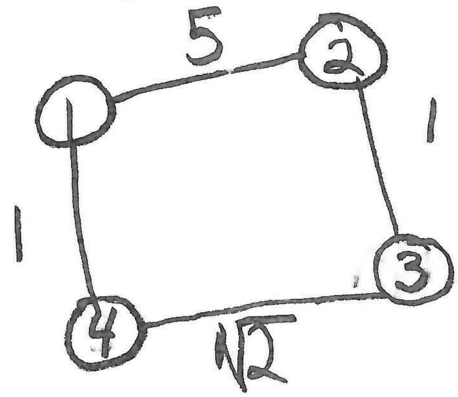
So, to have an EDM completion, every fully specified principal submatrix of a partial matrix must itself be an EDM , ie

a "partial EDM"

This is **not** always sufficient.

Example. For any cycle of 4 or more edges, positive weights on the edges give a partial EDM

e.g.



$$\begin{bmatrix} 0 & 25 & ? & 1 \\ 25 & 0 & 1 & ? \\ ? & 1 & 0 & 2 \\ 1 & ? & 2 & 0 \end{bmatrix}$$

each specified submatrix is 2-by-2, e.g.

$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$

But, there can be no EDM completion.

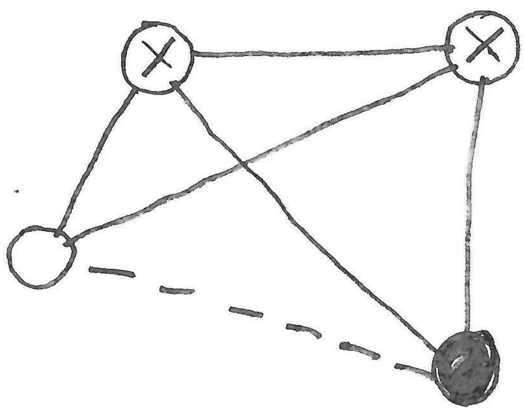
5 is too big!

Similary, for no graph with a long (≥ 4 edges) chordless cycle, can partial EDM be sufficient for an EDM completion.

"Bad" data for the cycle may always be embedded in partial EDM data for the entire graph.

So, what are the graphs for which partial EDM is sufficient for EDM completion?

Chordality is necessary.
It is also sufficient!



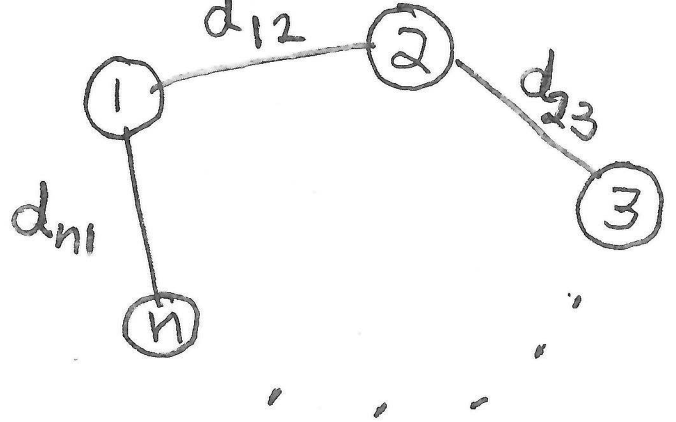
For any realization of the ⊗, ● triangle and the ⊗, ○ triangle the Euclidean length of the dashed edge is determined.
(May also see this analytically)

Since we may complete a chordal graph to a complete one by adding edges and maintaining chordality, the overlapping clique case is sufficient.

<u>Theorem</u>. Given a graph G, every partial EDM with graph G has an EDM completion iff G is chordal.

What about other graphs? More conditions are necessary.

The case of (simple) cycles



Cycle inequality suffices.

"maximum edge distance $\leq$ sum of the other distances"

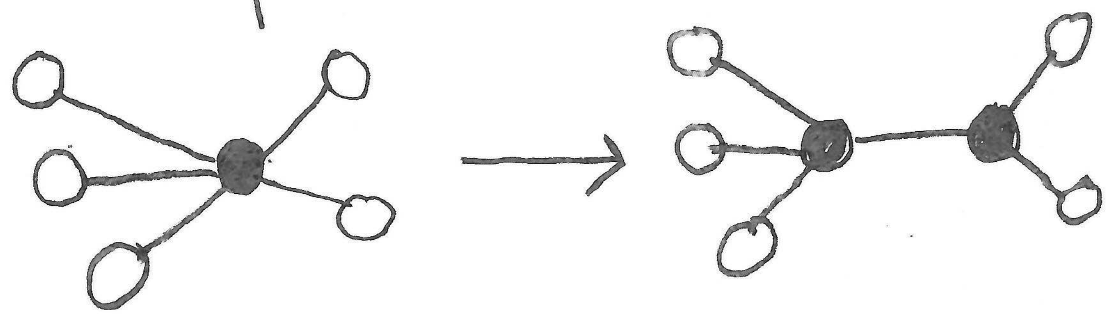
This condition is necessary on the induced cycles of any non-chordal graph.

When is it sufficient?

EDM is RPD (semi-algebraic).
So, for any graph, the EDM completable partial matrices may be identified by finitely many polynomial conditions.
(Hard to know what they are for a given graph.)

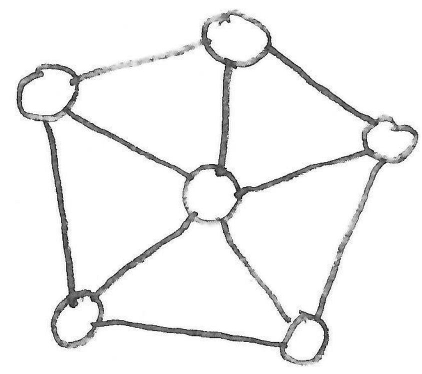
For which graphs are they the cycle conditions (+ partial EDM)?

Vertex partition



Graph G_2 is "built from" G_1 if G_2 may be obtained from a finite number (at least 1) of vertex partitions.

W_k : wheel on k vertices



W_6

Theorem. Every partial matrix, the graph of whose specified entries is G, and whose induced cycles satisfy the cycle inequalities has an EDM completion <u>iff</u>

G has no induced subgraph that is W_k, $k \geq 5$, nor that can be built from W_k, $k \geq 4$.

Notes: • other equivalent graph theoretical conditions
• can be checked in linear (?) time

Other graphs?

Selected References

M. Bakonyi and C.R. Johnson, The Euclidean distance matrix completion problem, *SIAM J. Matrix Anal. Appl.*, 16, No. 2, (1995), pp. 646-654.

L. M. Blumenthal, *Theory and Applications of Distance Geometry*, Clarendon Press, Oxford, 1953.

M. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, NY, 1980.

J. C. Gower, Properties of Euclidean and non-Euclidean distance matrices, , *Linear Algebra and Its Applications*, 67 (1985), pp. 81-97.

R. Grone, C. R. Johnson, E. Sa, and H. Wolkowicz, Positive definite completions of partial Hermitian matrices, *Linear Algebra Appl.*, 58 (1984), pp. 109-124.

C. R. Johnson, C. A. Jones, and B. K. Kroshchel, The Euclidean distance completion problem: Cycle completability, *Linear and Multilinear Algebra*, 39 (1995), pp. 195-207.

C. R. Johnson and P. Tarazaga, Connections between the real positive semidefinite and distance matrix completion problems, *Linear Algebra and Its Applications*, 223/224 (1995) pp. 375-391.

I. J. Schoenberg, Remarks to Maurice Frechet's article "Sur la definition axiomatique d'une classe d'espaces vectoriels distancies and applicables vectoriellement sur l'espace de Hilbert," *Ann. Math*. 36:724-732 (1935).