Beyond Euclidean Distance Matrices

Cem Ates Musluoglu, Adam James Scholefield

LCAV - School of Computer and Communication Sciences - EPFL - Lausanne, Switzerland

1. Problem statement

Study of the properties of matrices containing the distance between points $\mathbf{x}_i$ and linear varieties $\mathcal{F}_j = \left\{\mathbf{f} : \mathbf{f} = \mathbf{p}_j + \sum_{\ell=1}^{k} \mathbf{v}_j^\ell t_\ell\right\}$ of dimension k and reconstruction of the configuration based on distance measurements.

Key terms: Euclidean distance, linear varieties

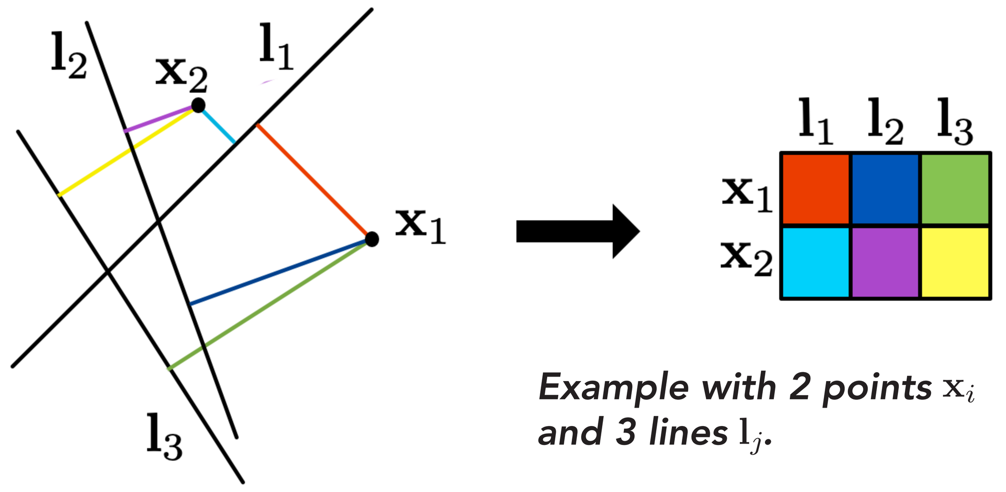
2. Distance matrix

For linear varieties $\mathcal{F}_j$, we assume $\mathbf{p}_j^T \mathbf{v}_j^\ell = 0$ and $||\mathbf{v}_j^\ell||^2 = 1, 1 \leq \ell \leq k$. The chosen metric to compute the distance between two linear varieties $\mathcal{F}$ and $\mathcal{G}$ given by the parametric pairs $(\mathbf{p}, \Phi)$ and $(\mathbf{q}, \Psi)$ is:

$$d(\mathcal{F}, \mathcal{G}) = \min_{\mathbf{f} \in \mathcal{F}, \mathbf{g} \in \mathcal{G}} ||\mathbf{f} - \mathbf{g}||^2 = ||C(\mathbf{p} - \mathbf{q})||^2,$$

where rows of C form an orthonormal basis for $\mathcal{N}(\Phi^T) \cap \mathcal{N}(\Psi^T)$. This metric makes the distances invariant to rigid motion of the whole configuration.

Symbol	Notation	Corresp. Matrix
$\mathbf{x}_i$	Point i	$X \in \mathbb{R}^{d \times n}$
$\mathcal{F}_j$	Linear variety j	-
$\mathbf{p}_j$	Intercept of linear variety j	$P \in \mathbb{R}^{d \times m}$
$\mathbf{v}_j^\ell$	ℓ-th direction vector for linear variety j	$V_\ell \in \mathbb{R}^{d \times m}$
d_{ij}	Distance between point i and linear variety j	$D \in \mathbb{R}^{n \times m}$



Example with 2 points $\mathbf{x}_i$ and 3 lines $\mathbf{l}_j$.

In the special case of distance between points $\{\mathbf{x}_i\}_{1 \leq i \leq n}$ and linear varieties $\{\mathcal{F}_j\}_{1 \leq j \leq m}$, the distance matrix can be written as:

$$D = \text{edm}(X, P) - \sum_{\ell=1}^{k} (X^T V_\ell)^{\circ 2},$$

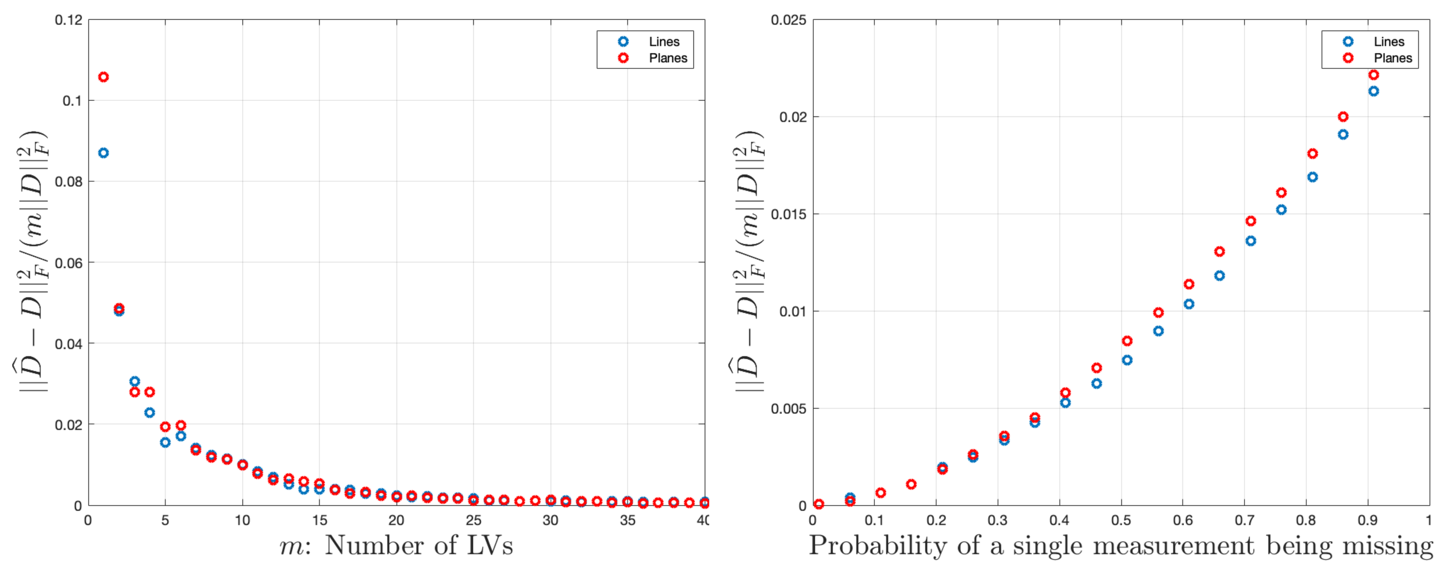
where $\text{edm}(X, P)$ contains the distances between the columns of X and the columns of P.

3. Completing the distances

The measurements are modeled by $\widetilde{D} = M \circ (D + Z)$, where M corresponds to a mask matrix putting to 0 missing measurements and Z represents the additive measurement noise. How to estimate the true values?

Theorem 1. $\text{rank}(D) \leq d + 1 + \binom{d+1}{2}$, *where d is the dimension of the ambient space.*

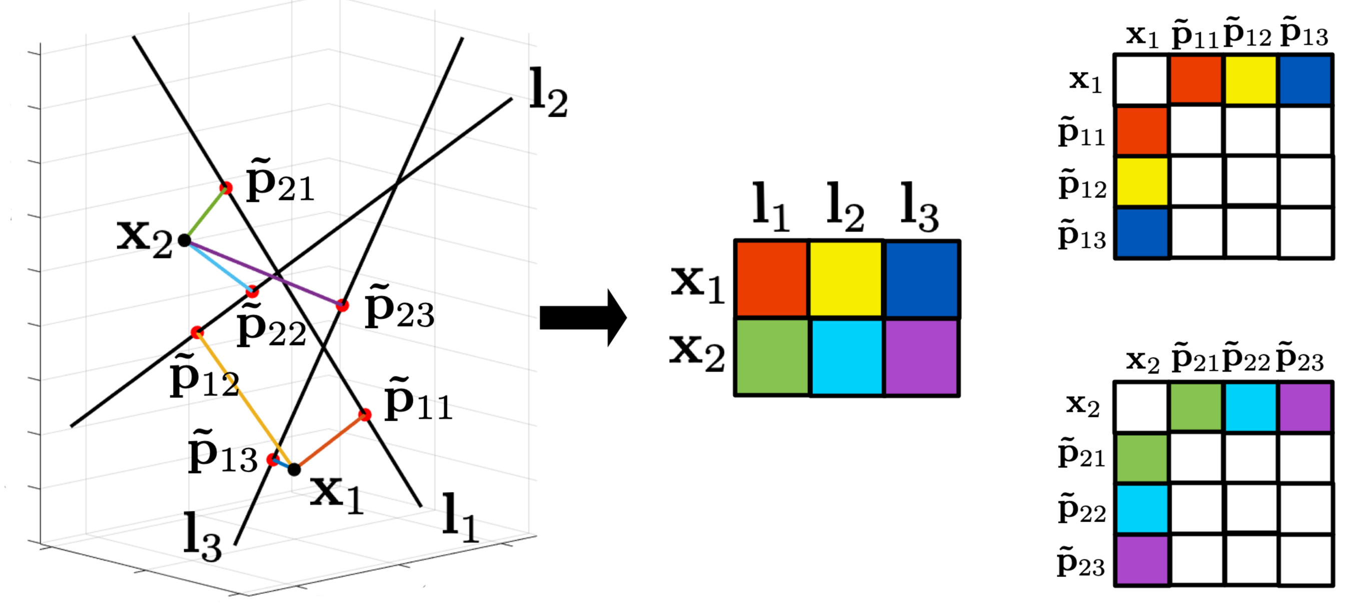
Proposed method: Applying the singular value decomposition to $\widetilde{D}$, we construct a low-rank approximation based on the result of **Theorem 1**: $\widehat{D} = \sum_{i=1}^{r} \widetilde{\sigma}_i \widetilde{\mathbf{u}}_i \widetilde{\mathbf{w}}_i$. The procedure is iterated by alternating between the low-rank approximation and forcing the known entries, i.e. where the mask is non-zero.



Estimation error for growing number of linear varieties (left) and probability of missing measurements (right).

4. Reconstructing the configuration

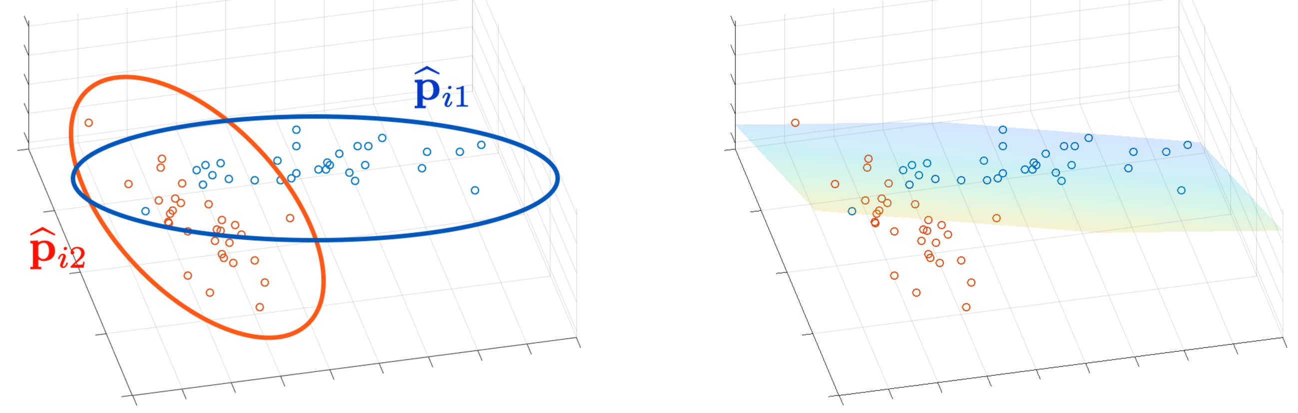
Proposed method: The point on linear variety j where the minimum distance with point i is obtained is $\widetilde{\mathbf{p}}_{ij} = \mathbf{p}_j + \Phi_j \Phi_j^T \mathbf{x}_i$, therefore, the distance matrix can be decomposed to n different EDMs, and the problem is translated to EDM completion.



Decomposition procedure.

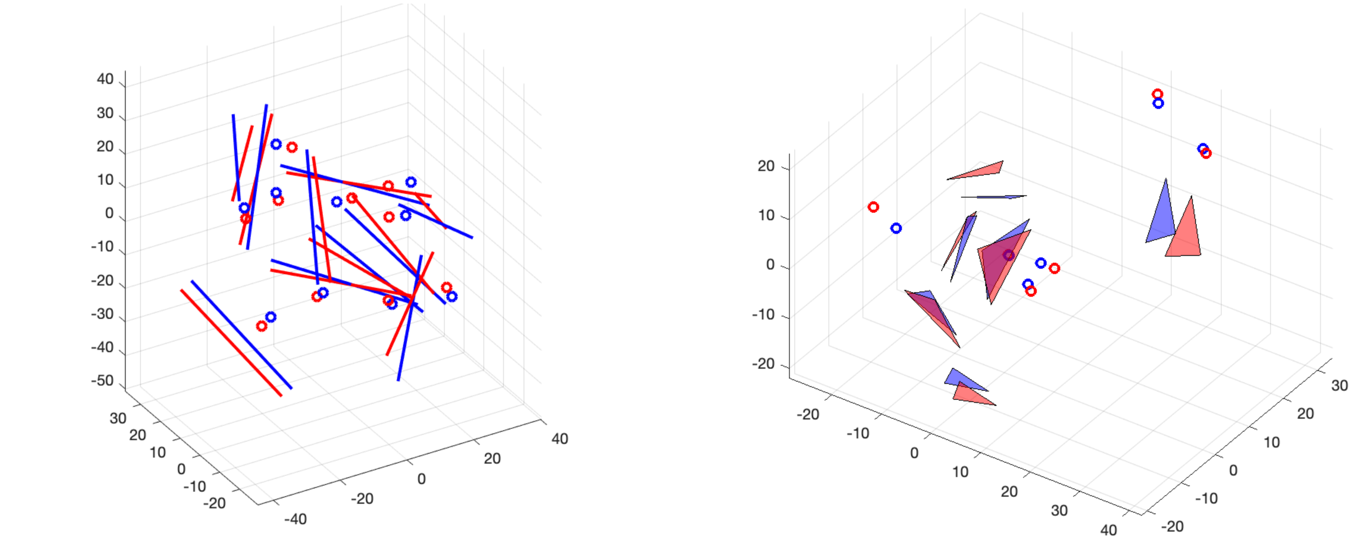
Step 1: Estimate the vectors $\widetilde{\mathbf{p}}_{ij}$ for a fixed i using EDM completion. The results are better when we know the distance between points $\mathbf{x}_i$.

Step 2: For a fixed j, every estimate $\widehat{\mathbf{p}}_{ij}$ of $\widetilde{\mathbf{p}}_{ij}$ should be on the same linear variety, therefore, we find the best fitting linear variety minimizing: $E_i\left[||N_j(\widehat{\mathbf{p}}_{ij} - \mathbf{p}_j^{\text{avg}})||^2\right]$, where the rows of N_j are the orthonormal vectors normal to the linear variety and $\mathbf{p}_j^{\text{avg}}$ is the average of the estimates over i.

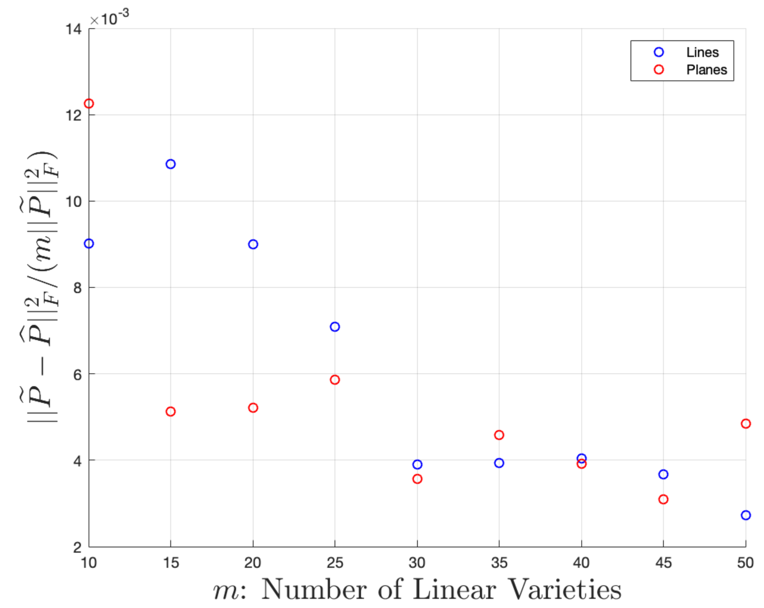


Example with 2 planes. For each cluster, we fit the plane minimizing the previous expectation.

5. Results



Example of reconstruction for lines and planes. Original configuration in red and estimation in blue.



Estimation error for growing number of linear varieties.

References:

[1] Ivan Dokmanic, Reza Parhizkar, Juri Ranieri, and Martin Vetterli. Euclidean distance matrices: essential theory, algorithms, and applications. IEEE Signal Processing Magazine, 32(6): 12–30, 2015.

[2] Miranda Krekovic, Ivan Dokmanic, and Martin Vetterli. Omnidirectional bats, point-to-plane distances, and the price of uniqueness. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 3261–3265. Ieee, 2017.

[3] Arthur M DuPré and Seymour Kass. Distance and parallelism between flats in rn. Linear Algebra and its Applications, 171:99–107, 1992.

email: cem.musluoglu@epfl.ch adam.scholefield@epfl.ch web: lcav.epfl.ch