COMP229: Introduction to Data Science
Lecture 22: Isometry invariants

Olga Anosova, O.Anosova@liverpool.ac.uk
Autumn 2023, Computer Science department
University of Liverpool, United Kingdom

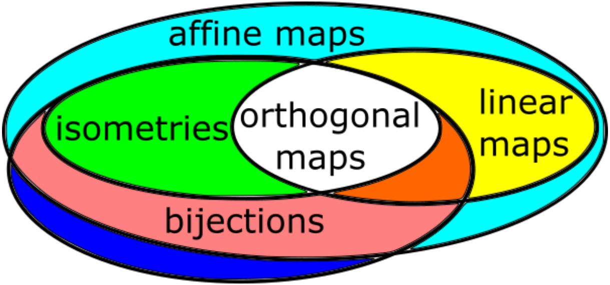
Lecture plan & learning outcomes

On this lecture we should learn

- that isometry is an equivalence.
- what is an invariant
- examples of isometric invariants for trianges.
- what is a complete invariant.
- that pairwise distances of point clouds are "good" invariants.

Reminder: orthogonal maps

- One of the important properties is our ability to see interconnections.



-

Isometry problem for clouds

Complicated rigid objects, e.g. mechanical parts, crystals, are often represented by a data cloud of points (corners, atoms) whose interpoint distances should be preserved under any equivalence.

Isometry problem : given two data clouds in $\mathbb{R}^m$, how can we decide if they are *isometric*, i.e. there is an isometry that maps one into another?

Is it possible that A is isometric to B, which is isometric to C, but C isn't isometric to A?

Isometries are equivalences

Claim 22.1. The isometries define an equivalence relation on point clouds (and other spaces), i.e.

the identity map $f(p) = p$ is an isometry;

the inverse f^{-1} of any isometry is an isometry;

any composition of isometries is an isometry.

Proof. All properties follow from the definition that an isometry preserves distances. The symmetry follows, because any isometry f is bijective. $\square$

How to classify up to isometry

For any equivalence relation, all objects split into non-overlapping classes so that only objects in the same class should be equivalent to each other, but any objects from different classes are not equivalent.

An isometry of $\mathbb{R}^2$ applied to triangles is called a *congruence*. How can we distinguish and classify triangles (3-point clouds) in $\mathbb{R}^2$?

SSS theorem is easier to work with than SAS or ASA.

Easy and hard parts of the problem

Problem 22.2. (from exam 2019) Are the triangles on the vertices below isometric or not? First: $(0,0),(4,0),(0,3)$; second: $(1,1),(5,1),(1,4)$.

Solution 22.2. Though the vertices have different coordinates: $(0,0),(4,0),(0,3)$; $(1,1),(5,1),(1,4)$, we can't claim that triangles are not isometric.

The triangles are isometric, because the second is the first triangle translated by $\vec{u}=(1,1)^T$. $\square$

The easy part of a classification is to show an equivalence by getting one object from another.

The harder part: show that objects are different, why can't one of infinitely many equivalences match them?

Invariants help distinguish objects

Definition 22.3. An **invariant** of objects considered up to an equivalence relation is a function $\mathcal{I}$ that takes the *same value on all equivalent objects*.

Invariant $\mathcal{I}$: A is equivalent to $B \Rightarrow \mathcal{I}(A) = \mathcal{I}(B)$.

$$\mathcal{I} : \frac{\text{objects}}{\text{equivalence}} = \left(\begin{array}{c} \text{classes of} \\ \text{equivalence} \end{array} \right) \to \begin{array}{c} \text{simple} \\ \text{values} \end{array}$$

Example: The number of points is an invariant of clouds.

Claim 22.4. If an invariant takes different values on two objects, then these objects are different (non-equivalent). $\square$

MU game

Suppose 3 symbols **M, I**, and **U** can be combined to produce strings. Start with the string **MI** and transform it into the string **MU** using the following rules:

1. $xI \rightarrow xIU$ Add a U to the end of any string ending in I: $MI \rightarrow MIU$
2. $Mx \rightarrow Mxx$ Double the string after the M: $MIU \rightarrow MIUIU$
3. $xIIIy \rightarrow xUy$ Replace any III with a U: $MIIIU \rightarrow MUU$
4. $xUUy \rightarrow xy$ Remove any UU: $MUUU \rightarrow MU$

What is the minimal set of transformations that changes **MI** into **MU**?

MU invariant

$\mathcal{I}=$ *"The number of I's in the string is not a multiple of 3."*

Check that $\mathcal{I}$ is an invariant:

1. In $xI \rightarrow xIU$ the number of I's doesn't change, $\mathcal{I} = \mathcal{I}$.
2. In $Mx \rightarrow Mxx$ the number of I's doubles, divisibility by 3 is the same and $\mathcal{I} = \mathcal{I}$.
3. In $xIIIy \rightarrow xUy$ the number of I's decreases by 3, $\mathcal{I} = \mathcal{I}$.
4. In $xUUy \rightarrow xy$ the number of I's is unchanged, $\mathcal{I} = \mathcal{I}$.

Appy to our problem:

the number of I's in MI

MU invariant

$\mathcal{I}=$ *"The number of I's in the string is not a multiple of 3."*

Check that $\mathcal{I}$ is an invariant:

1. In $xI \to xIU$ the number of I's doesn't change, $\mathcal{I} = \mathcal{I}$.
2. In $Mx \to Mxx$ the number of I's doubles, divisibility by 3 is the same and $\mathcal{I} = \mathcal{I}$.
3. In $xIIIy \to xUy$ the number of I's decreases by 3, $\mathcal{I} = \mathcal{I}$.
4. In $xUUy \to xy$ the number of I's is unchanged, $\mathcal{I} = \mathcal{I}$.

Appy to our problem:

the number of I's in MI is equal to 1, and in MU this number is 0, hence the task is impossible.

無

Out of the system

An algorithm can generate every valid string of symbols, and would search forever, never seeing the futility.

A human player will begin to suspect that the puzzle may be impossible. Then one "jumps out of the system" and starts to reason *about* the system instead of *within*.

Eventually, one realises that the system is in some way about a completely different issue of *divisibility by three*.

On this outer level, the MU puzzle can be seen to be impossible.

There is currently no general automated tool that can detect this invariant, but once the invariant is introduced, a computer easily checks the rest.

Example invariants

A typical mistake is to classify objects by using non-invariants, e.g. people in photos by the colour of their clothes.

For triangles in $\mathbb{R}^m$: *non-invariants* (under all isometries) are

- positions of vertices,
- a barycentre,

invariants are

- lengths,
- angles,
- area.

If an invariant takes the same value on two objects, what can we conclude? Nothing! The height of a person is the invariant, millions have equal heights.

Invariants vs complete invariants

Definition 22.5. An invariant $\mathcal{I}$ is **complete** if $\mathcal{I}$ takes the same value only on equivalent objects.

Complete I: $\mathcal{I}(A) = \mathcal{I}(B) \Rightarrow A$ is equivalent to B.

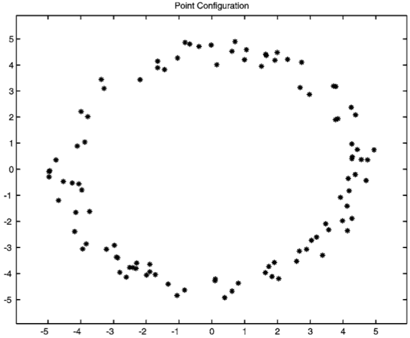
Are the following measurements complete human invariants: fingerprints, DNA?

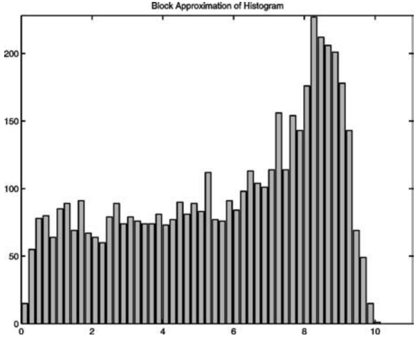
Claim 22.6. For triangles (3-point clouds), a complete invariant consists of 3 pairwise distances.

Proof. Match longest edges of equal lengths. Then then 3rd vertices of the two triangles coincide or are related by the reflection over the longest side. $\square$

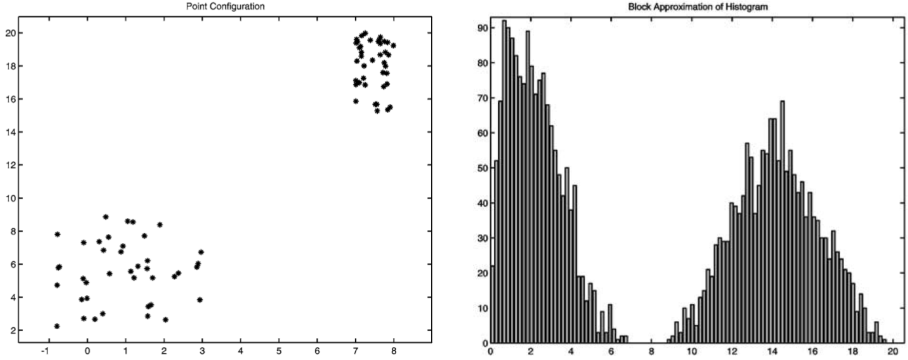
Circular cloud and its distribution

The histogram on the right has vertical bars. The height of each bar is the number of pairwise distances that fall within a short interval (bin). The histogram contains distances of all lengths from short to long.





2-cluster cloud and its distribution

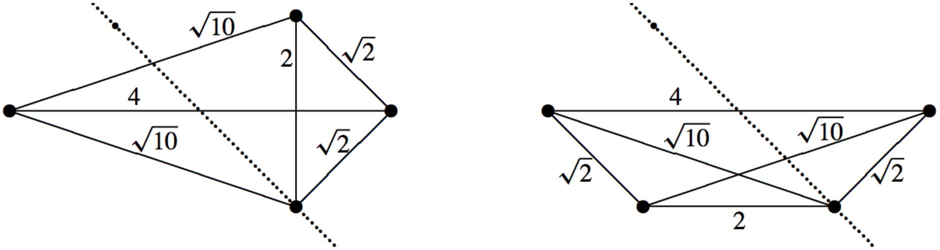


The histogram on the right contains many short distances (within clusters), many long distances (between clusters) and few mid-range distances.

Hence pairwise distance distributions can be used for comparing clouds in *general position* in $\mathbb{R}^n$, let's see why.

Interesting 4-point clouds

Example 22.7. The 4-point clouds below have the same distribution of 6 pairwise distances: $\sqrt{2}$, $\sqrt{2}$, 2, $\sqrt{10}$, $\sqrt{10}$, 4, but are not isometric, because their quadrilaterals have different areas. The distribution is not a complete isometry invariant of clouds.



All distances in one polynomial

There are larger non-isometric clouds with the same distribution of distances. These clouds cannot be uniquely reconstructed from all pairwise distances.

All distances in one polynomial

There are larger non-isometric clouds with the same distribution of distances. These clouds cannot be uniquely reconstructed from all pairwise distances.

But in which cases pairwise distances is a complete invariant?

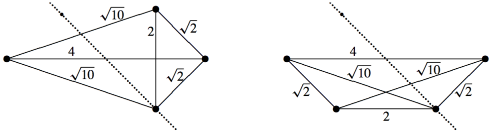
How to compare all pairwise distances? Is a matrix of all pairwise distances a good choice? No, because a matrix relies on the order.

Definition 22.8. Label points in a cloud C by $1, 2, \ldots, n$. Let d_{ij} be the distance between the i-th and j-th points. The **distance polynomial** is $F_C(x) = \prod\limits_{1 \leqslant i < j \leqslant n} (x - d_{ij})$, the product of linear factors $x - d_{ij}$, where x is a real variable.

For 3 points $F_C(x) = (x - d_{12})(x - d_{23})(x - d_{13})$.

Reconstructible configurations

Definition 22.9. A cloud C is **reconstructible from distances** if for any other cloud C' with the same distance polynomial ($F_C(x) = F_{C'}(x)$ for all x) there is an isometry of $\mathbb{R}^m$ that maps C to C'.



Those 4-point clouds C, C' are not reconstructible from distances, because C, C' are not isometric, but the distance polynomials are equal:

$$F_C = (x - \sqrt{2})^2(x - 2)(x - \sqrt{10})^2(x - 4) = F_{C'}.$$

Luckily these are 'almost all' exceptions. Why?

Reconstructible configurations

Theorem 22.10. (no proof needed for the exam)
For any $n \geqslant m + 2$, there is a polynomial $\mathcal{F}(C)$ depending on (all coordinates of) n points of a cloud $C \subset \mathbb{R}^m$ such that if $\mathcal{F}(C) \neq 0$, then the cloud C is reconstructible from distances.

General fact: For any non-zero polynomial $\mathcal{G}$ depending on mn coordinates of n points from C, a random cloud C satisfies $\mathcal{G}(C) \neq 0$ with a high probability.

Hence 'almost any' C is reconstructible from distances.

Time to revise and ask questions

- *Invariant* $\mathcal{I}$: A is equivalent to $B \Rightarrow \mathcal{I}(A) = \mathcal{I}(B)$.
- *Complete* $\mathcal{I}$: $\mathcal{I}(A) = \mathcal{I}(B) \Rightarrow A$ is equivalent to B.
- The distribution of all pairwise distances is an isometry invariant of clouds, 'almost' complete.

Problem 22.11. Is the average distance between points an isometry invariant, a complete invariant? Why or why not?

References & links

- On reconstructing n-point configurations from the distribution of distances or areas. https://doi.org/10.1016/S0196-8858(03)00101-5
- MU puzzle
- Gödel's incompleteness theorems about the inevitability of breaking the system for new discoveries.