COMP229: Introduction to Data Science
Lecture 5: isometric point clouds

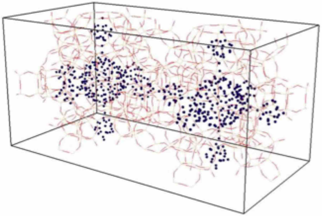
Vitaliy Kurlin, vitaliy.kurlin@liverpool.ac.uk
Autumn 2018, Computer Science department
University of Liverpool, United Kingdom

Use your card with the scanner for registration

Point clouds in real data

A real object is often represented (or described) by a cloud of points, for example

- *edgels* (edge pixels where intensity suddenly changes) are simple descriptors of images
- noisy (low quality) scans of visual markers
- point samples on surfaces around molecules



Cloud classification problem

Example problem. Decide if any two clouds in $\mathbb{R}^m$ (containing the same number of points) are isometric, i.e. can be obtained from each other by translations, rotations, reflections.

For 3-point clouds (triangles, also degenerate ones with 3 vertices in the same line), a complete invariant consists of 3 edge-lengths by Claim 4.8.

How many pairwise distances between n points (distinct triangles on n points) can we count?

How many pairwise distances?

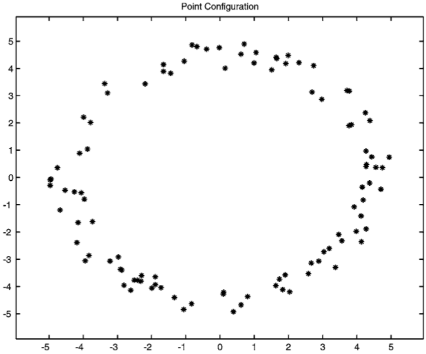
Claim 5.2. For n distinct points in any metric space, there are $\dfrac{n(n-1)}{2}$ pairwise distances, which can be computed in time $O(n^2)$, i.e. we need at most Bn^2 operations for a constant B.

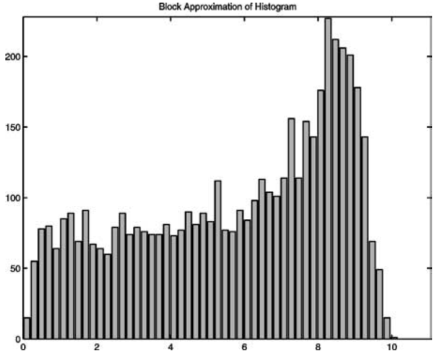
Proof. Any of n points has $n-1$ distances to the remaining points. By symmetry these $n(n-1)$ distances split into $\dfrac{n(n-1)}{2}$ pairs. Each distance needs only a constant number of operations. $\square$

Can all distances distinguish 4-point clouds?

Circular cloud and its distribution

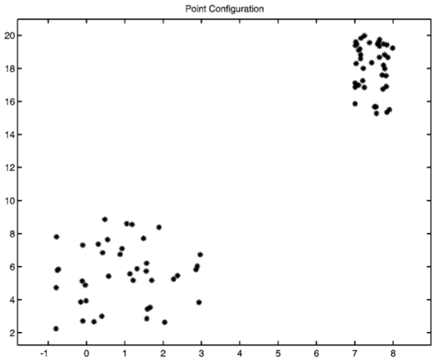
The histogram on the right contains vertical bars.
Each bar counts distances within a short range.

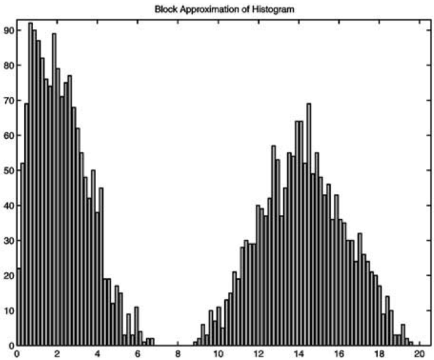




The histogram contains distances of all lengths
from short to long, more longer distances.

2-cluster cloud and its distribution

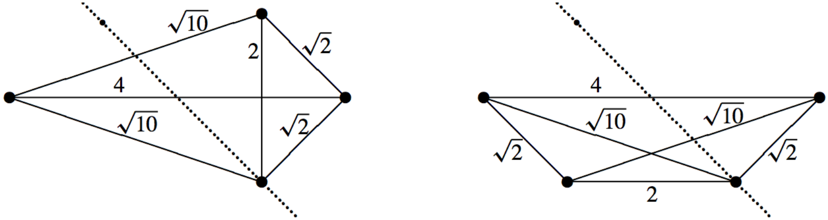




The histogram contains many short distances (within clusters), many long distances (between clusters) and very few mid-range distances.

Interesting 4-point clouds

The 4-point clouds below have the exactly same distribution of 6 pairwise distances: $\sqrt{2}$, $\sqrt{2}$, 2, $\sqrt{10}$, $\sqrt{10}$, 4. Add the 5th marked point to each cloud to see how these clouds are related.



The 4-point clouds are not isometric, because their quadrilaterals have different areas.

All distances in one polynomial

Add more points at corresponding positions in the dotted line gives larger clouds with the same distribution of distances. In general, a cloud cannot be uniquely reconstructed from distances.

Definition 5.2. 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 $F_C(x) = \prod_{1\le i<j\le n} (x - d_{ij})$, the product of linear factors $x - d_{ij}$, where x is a real variable, e.g. $F_C(x) = (x - d_{12})(x - d_{23})(x - d_{13})$ for 3 points.

Reconstructible configurations

Definition 5.3. A cloud C is *reconstructible from distances* if for any other cloud C' such that their distance polynomials are equal ($F_C(x) = F_{C'}(x)$ for all x) there is an isometry of $\mathbb{R}^m$ sending C to C'.

The 4-point clouds C, C' from slide 7 are not reconstructible from distances, because C, C' are not isometric, but their polynomials are equal:
$F_C(x) = (x-\sqrt{2})^2(x-2)(x-\sqrt{10})^2(x-4) = F_{C'}(x)$.
Luckily these are "almost all" exceptions.

Reconstructible configurations

Claim 5.4. [no proof needed, Theorem 2.6 in Boutin, Kemper, Advances in Appl. Maths (2004)]

For any $n \geq m + 2$, there is a non-zero polynomial $f(C)$ depending on (all coordinates of) n points of cloud $C \subset \mathbb{R}^m$ such that if $f(C) \neq 0$ then the cloud C is reconstructible from distances.

For any non-zero polynomial f depending on mn coordinates of n points from C, a random cloud C satisfies $f(C) \neq 0$ with a high probability. Briefly, "almost any" C is reconstructible from distances.

Your questions and the quiz

To benefit from the lecture, now you could

- ask or submit your anonymous questions to the COMP229 folder after the lecture;
- write down your summary in 2–3 phrases, e.g. list key concepts you have learned;
- talk to your classmates to revise the lecture.

Question. How many distinct triangles on n points can we count?

Answer to the quiz and summary

Answer. We have n choices for a 1st vertex, $n-1$ for a 2nd vertex, $n-2$ for a 3rd vertex. The resulting $n(n-1)(n-2)$ triangles split into 6-tuples related by 6 permutations of 3 vertices.

- $\dfrac{n(n-1)}{2}$ pairwise distances between n points.
- $\dfrac{n(n-1)(n-2)}{6}$ distinct triangles on n points.
- "Almost all" clouds in $\mathbb{R}^m$ are reconstructible from the distribution of pairwise distances.