Metrics and network cosntruction

Social and Technological Networks

Rik Sarkar

University of Edinburgh, 2019.

Metric

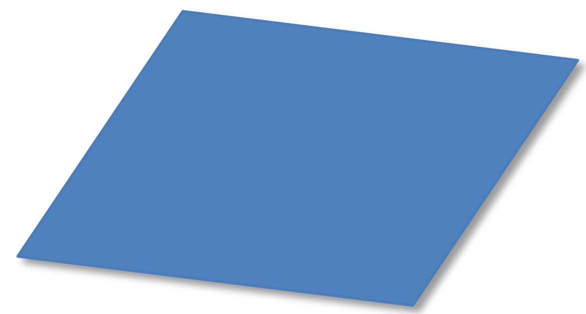
- A distance measure d on a set X
- Satisfies usual intuitions, triangle inequality

Metrics

- Metrics are Important because:
  - Metrics are used to construct networks
  - Networks have metrics that determine their properties

Euclidean metric

- 1-D
  - Straight line (think x-axis)



- 2-D
  - Plane
- Distance measure in dimension d:

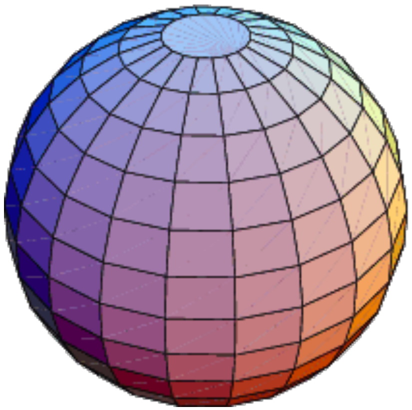
$$d(u, v) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_d - v_d)^2}$$

Non Euclidean metrics

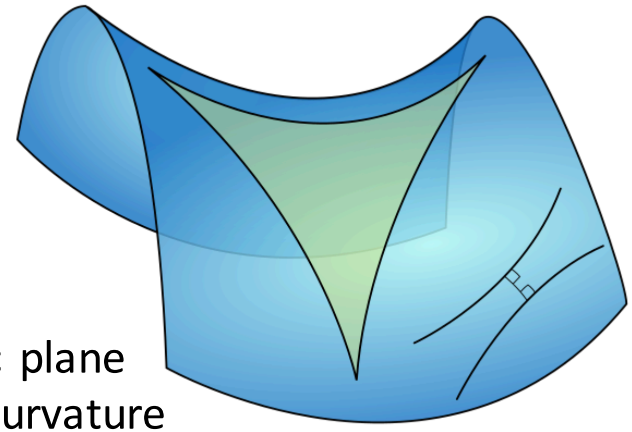
- A lot of maths for Euclidean metrics

- What are examples of non-Euclidean metrics?

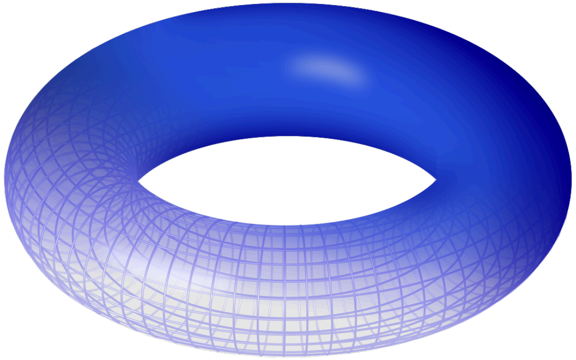
Non-euclidean metrics



Sphere
Positive curvature



Hyperbolic plane
Negative curvature



Realistic shapes. With bends, and cycles.

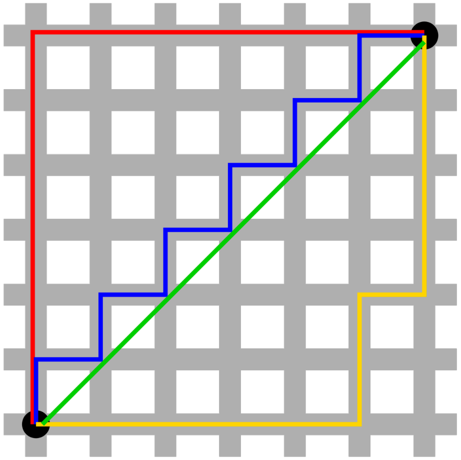
- L_p metrics

$$d(u,v) = \sqrt[p]{(u_1 - v_1)^p + (u_2 - v_2)^p + \cdots + (u_d - v_d)^p}$$

L_1 metric

- Manhattan distances

$$d(u,v) = |u_x - v_x| + |u_y - v_y|$$



L_∞ Metric

- Largest component over dimensions

$$d(u,v) = \lim_{p \to \infty} \sqrt[p]{(u_x - v_x)^p + (u_y - v_y)^p}$$

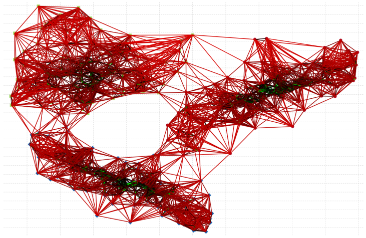
$$d(u,v) = \max(|u_x - v_x|, |u_y - v_y|)$$

The undirected shortest path distance

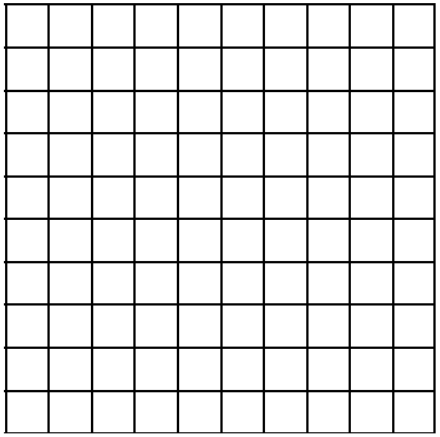
- Is a metric
- In unweighted graphs, all values are integers

Graph Embedding

- Map the vertices V to points in the plane
  - (or some other space)

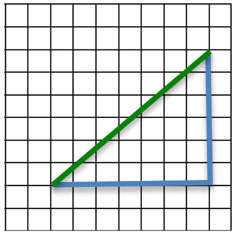


- Usually, different vertices are mapped to different points

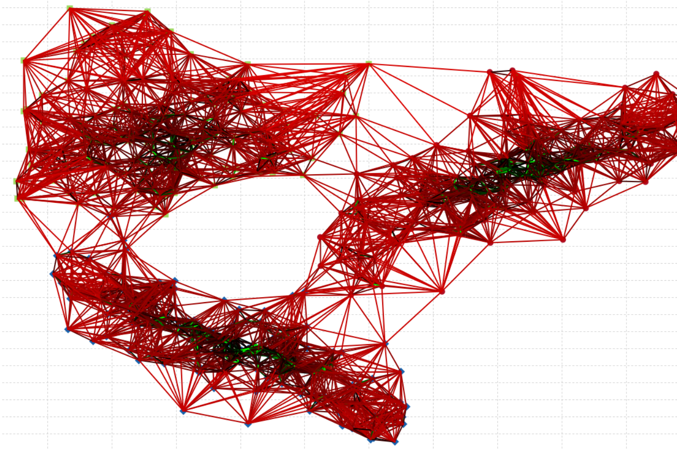


Different distances

- What is the distance between u and v?
- Possibility 1 (Embedding or extrinsic distance):
  - Distance in the embedded space
    - E.g. Euclidean distance



- Possibility 2 (Intrinsic distance):
  - Distance in the graph
    - The length of shortest path
- Possibility 3 (Intrinsic distance):
  - Weighted distance in the graph
    - Weight/length given by embedding



Where do metrics come from?

- Possibility 1:
  - We are given weights/lengths of edges
- Possibility 2a:
  - Vertex locations are given. Eg. Mobile phone locations
- Possibility 2b:
  - We are given real valued features like age, salary, etc
  - We can use these as dimensions and compute distances.

- We will often use and compare multiple metrics on the same network

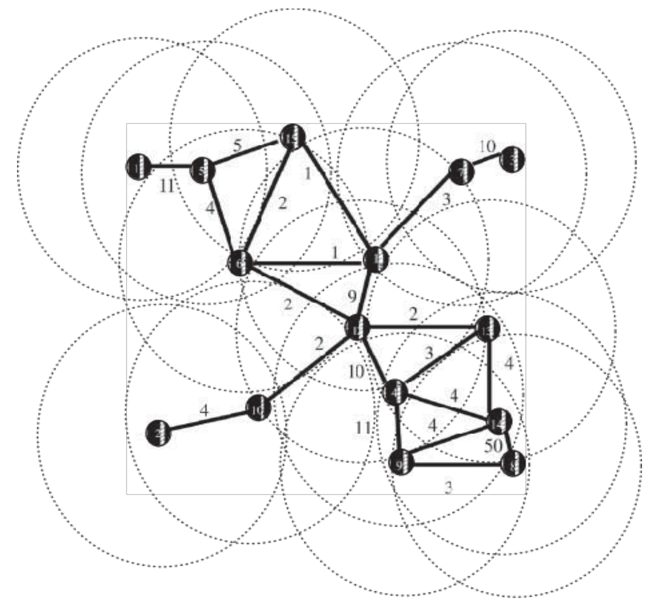
- E.g. on a map
  - The Eulcidean distance between nodes (junctions)
  - The distance along road networks
  - The travel time at busy hours etc

- E.g. For people in a social network
  - The shortest path distance on the unweighted graph
  - The shortest path, where weights are given according to strength of friendship
  - Distance between nodes after an embedding in k-dimensional space
  - Distance after embedding by (age, salary, location)
  - Distance on the UDG or k-NN after embedding by some features…

- (Which of these are intrinsic and which are extrinsic?)

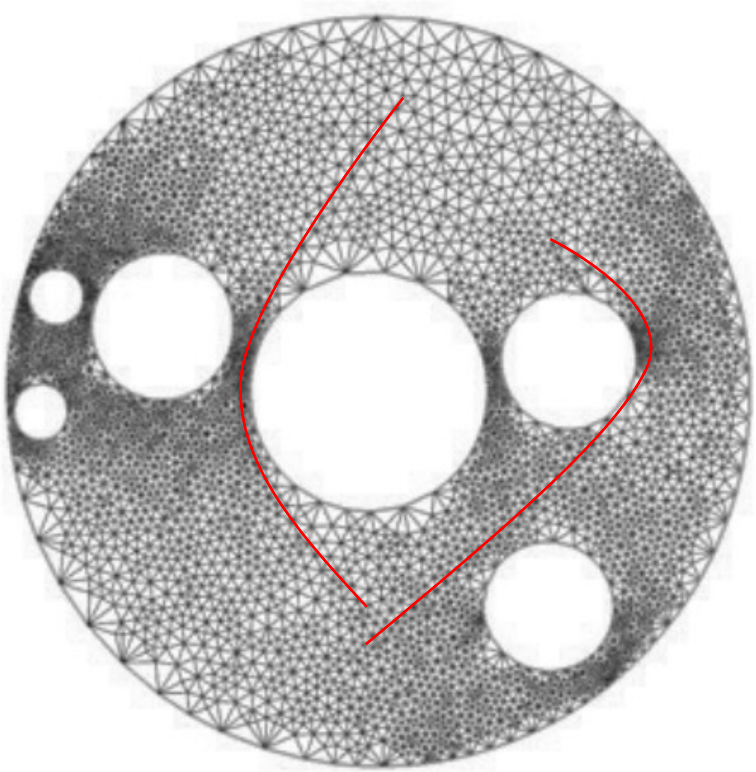
Making networks from metrics

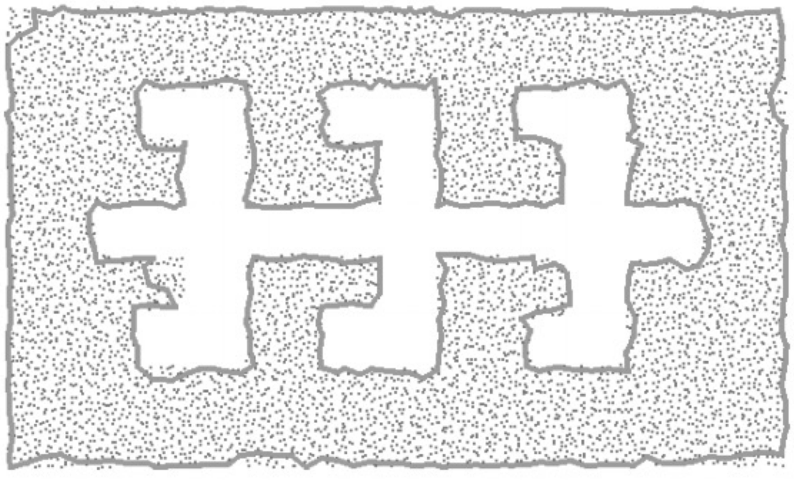
- Unit disk graphs:
  - Consider vertices in the plane (like wireless nodes)
  - Connect two vertices by an edge if they are within distance 1 of each-other. (within transmission distance)
  - Applies generally to higher dim (Unit ball graphs)
  - Connect two nodes if they are within a given distance



Network Metric: Shape of the data

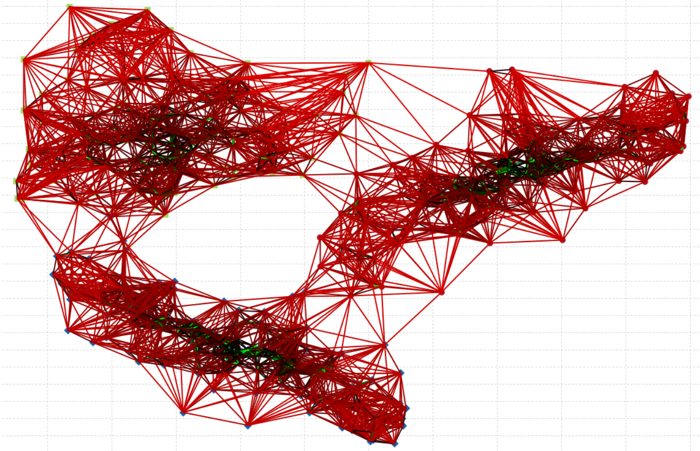
- Intrinsic metric determined by shortest paths





k-NN graphs

- For each vertex, find k nearest neighbors
  - Connect edges to all k nearest neighbors
  - Variants:
    - Connect all k-NN edges
    - Connect only if both vertices are k-NN of each-other



Network construction

- Given any dataset with distances between items, we can construct a network

Finding distance between two nodes in a graph

- Breadth first search (for unweighted graphs)

- Dijkstra’s shortest path algorithm (for weigted graphs)

- All pairs shortest paths
  - Floyd Warshall Algorithm