A metric space is a set together with a notion of distance b/w its elements, called points. The distance is measured by a function called a <u>metric</u> or <u>distance function</u>.

A metric space is an <u>ordered pair</u> (M, d), where M is a set and d is a metric on M, i.e. a function

$$d : M \times M \longrightarrow \mathbb{R}$$

satisfying the following axioms for all points $x, y, z \in M$:

1. $d(x, x) = 0$

2. $d(x, y) > 0$, if $x \neq y$

3. $d(x, y) = d(y, x)$ [symmetry]

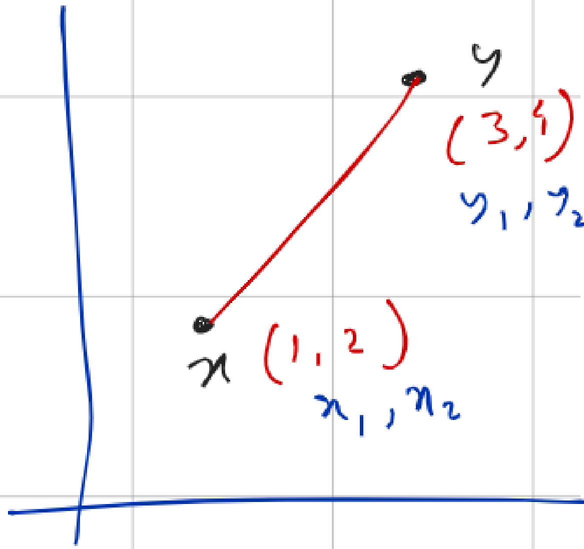
4. Triangle inequality: $d(x, y) + d(y, z) \geq d(x, z)$

Euclidean

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

$$\rightarrow = \sqrt{(1-3)^2 + (2-4)^2}$$

$$\rightarrow = \sqrt{8} = 2.82$$

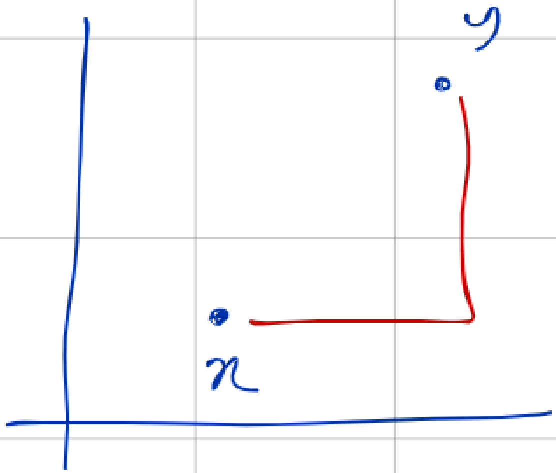


Manhattan / City Block / Taxi cab distance

$$d(x,y) = \sum_{i=1}^{n} |x_i - y_i|$$

$$\rightarrow = |1-3| + |2-4|$$

$$\rightarrow = 4$$



x:
$\{x_1, x_2, x_3, \ldots, x_n\}$

y:
$\{y_1, y_2, y_3, \ldots, y_n\}$

A NORM is a function from a real / complex vector space to the non negative real no.s that behaves in certain ways like the distance from the origin

The Euclidean distance in a Euclidean space is defined by a norm on the associated Euclidean vector space, called EUCLIDEAN NORM / L2 NORM / $\|d(x,y)\|_2$ or L^2

⊙ Think on: L1 - NORM

If interested, may check "QUATERNIONS"

Chebyshev Distance

$$d(x,y) = \max_i \left(|x_i - y_i| \right)$$

$$\max \left(|1-3|, |1-4| \right)$$

$$= \max(2, 3)$$

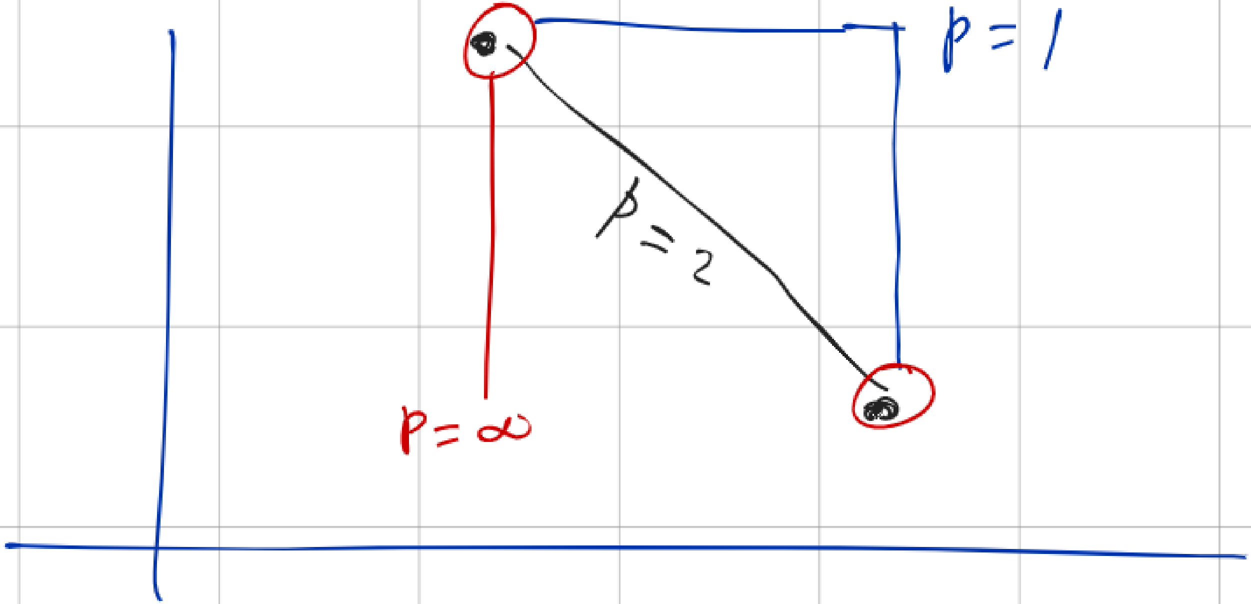
$$= 3$$

L_∞ metric

Think: The chebyshev distance is the limiting case of the order-p Minkowski, when p reaches infinity

MINKOWSKI DISTANCE

$$d(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$$



Cosine Similarity

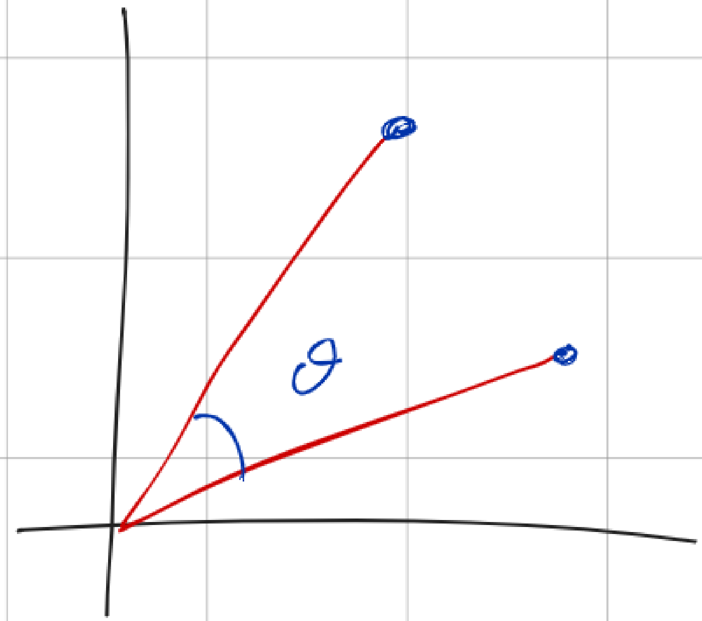
$[-1, 1]$

opposite orientation

same orientation

$$x \cdot y = \|x\| \, \|y\| \cos\theta$$

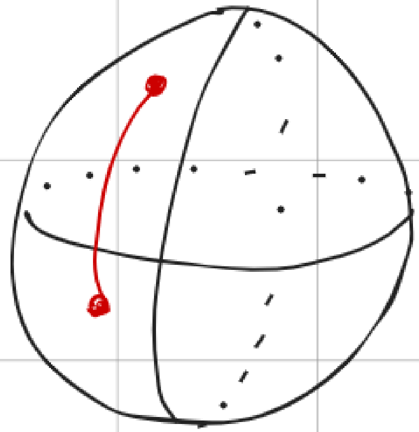
two non zero vectors



$$d(x,y) = \cos\theta = \frac{x \cdot y}{\|x\| \, \|y\|}$$

HAVERSINE distance

e.g. distance b/w 2 points on a sphere



formula : try to find by yourself

HAMMING distance

0 0 1 1 0 1

0 0 1 0 0 0

2

0 0 0 0

1 1 1 1

?

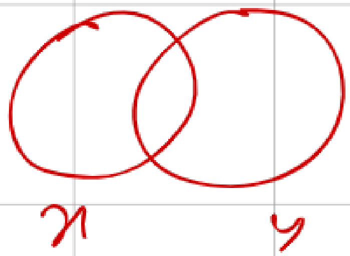
C H A N D R A

C H O N D R O

2

Jaccard index / Intersection over Union (IOU)

$$IOU = \frac{|x \cap y|}{|x \cup y|} = J(x,y)$$



$$0 \le J(x,y) \le 1$$

$|x \cap y| = 0$

$|x \cap y| = |x \cup y| = x = y$

JACCARD DISTANCE

Symmetric diff.

$$d_J(x,y) = 1 - J(x,y)$$

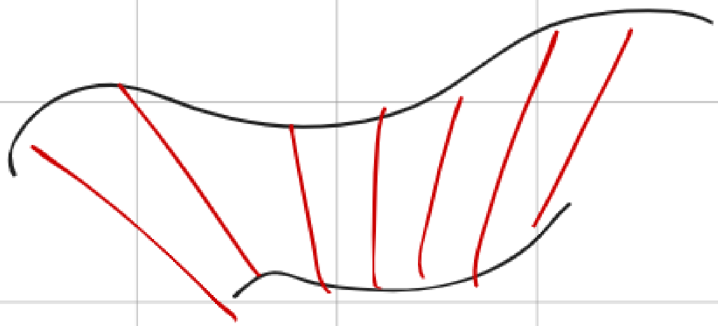
$$= \frac{|x \cup y| - |x \cap y|}{|x \cup y|} = \frac{x \Delta y}{|x \cup y|}$$

Check : SÖRENSEN DICE INDEX

Self Study : DTW (Dynamic Time Warping)

(If interested)

measures the distance b/w 2 time series of diff. length



MAHALANOBIS DISTANCE (MD)

It is a measure of the distance b/w a point $\underline{x}$ and a prob. dist. $\mathcal{G}$.

prob. dist. $\mathcal{G}$ on $\mathbb{R}^N$ with mean $\underline{\mu} = (\mu_1, \mu_2, \cdots, \mu_N)^T$

Positive semi definite COV. MATRIX Σ

MD of a point $\underline{x} = (x_1, x_2, \ldots, x_n)^T$ from $\mathcal{G}$ is

$$d_M(\underline{x}, \mathcal{G}) = \sqrt{(\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})}$$

(Since Σ is positive semi definite, so is Σ^{-1}, thus the square roots are always defined.)

Check BHATTACHARYA distance

(if interested)