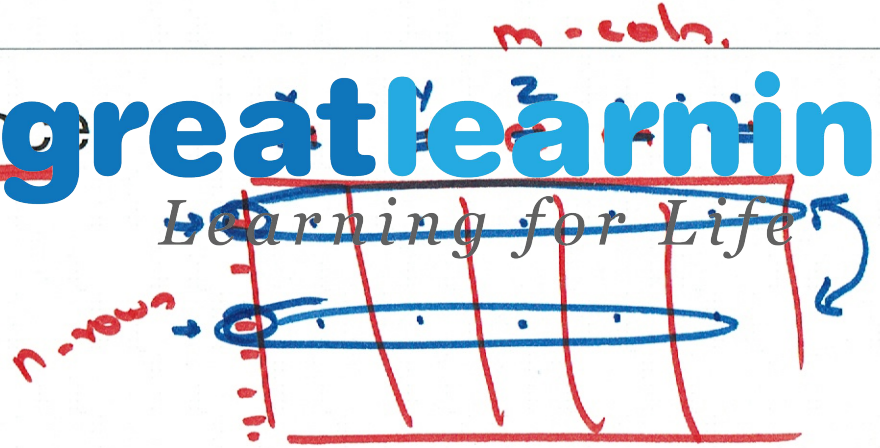
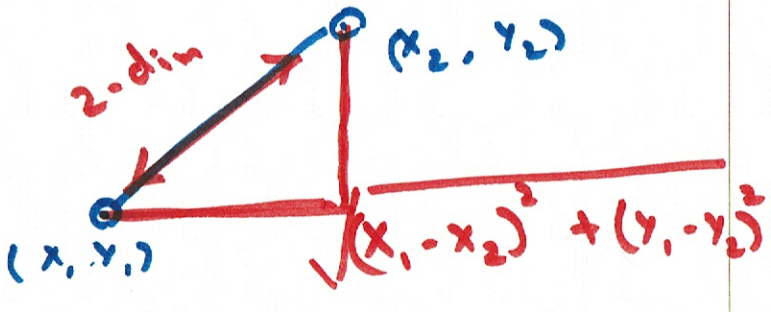
Distance



- Do define "similarity" you need a measure of distance
- Examples of common distance measures
  - Manhattan Distance
  - Eucledian Distance
  - Chebyshev Distance



$|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2| + \ldots$

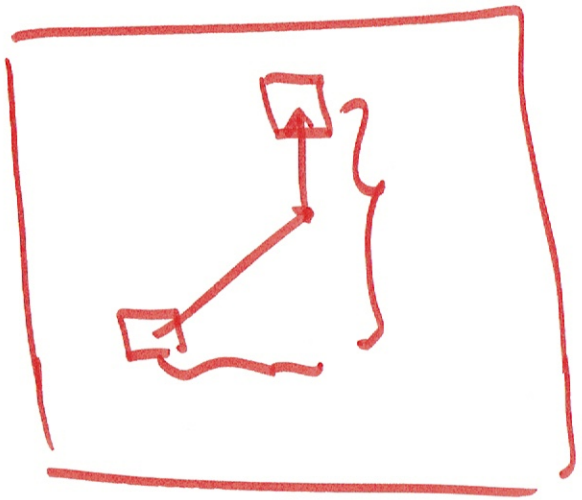
m-dim

$\sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 + \ldots]}$

Chebyshev (or) the Chessboard dist.

m-dim

$$\max(|x_1 - x_2|, |y_1 - y_2|, |z_1 - z_2|, \ldots)$$



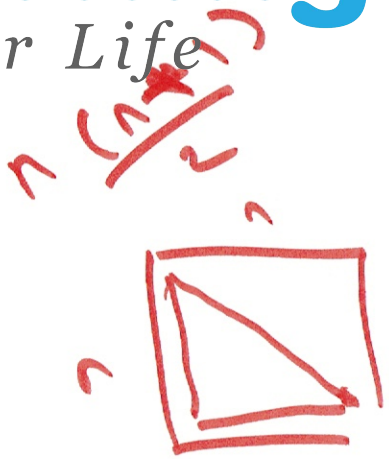
Minkowski

$$\left(\sum_{i=1}^{m} |x_i - y_i|^p \right)^{\frac{1}{p}}$$

if $p = 2$ ⇒ Euclidian dist.

if $p = 1$ ⇒ Manhattan dist.

if $p = \infty$ ⇒ Chebyshev dist.



Connecting based on roughly begin by computing
500,000 dist $\frac{n \times (n+1)}{2}$

Centroid based on roughly begin by computin
5 - group 5 x 1000 dist. } 5 x n