Module 2

Motivation

It's often necessary to measure similarity between data points.

In context of dimensionality reduction, we are interested in finding compact representations of data that live in lower dimensional space, but which is similar to original data.

One of the key concepts that we will be looking at in this case is orthogonality.

We use orthogonal projections of data points as a way to compress the data, while minimizing the loss of information.

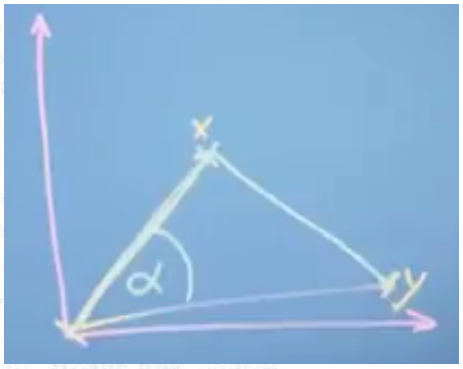
Thinking of data points as vectors in vector space, will allow us to look at two types of similarities between data points

① Distance of data points from each other.

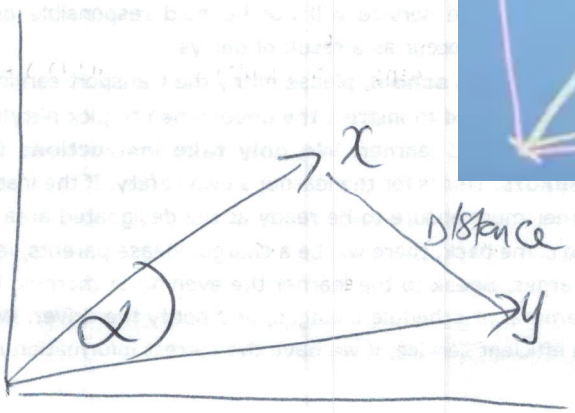
② The angle between the data points

Let's have look at an example:

Assume we have, we living in a two dimensional world, and we have 2 vectors, x, and y.



Please



Things that we will be discussing in this course are:

- how to compute the length of individual vectors, i.e length of x

- looking at angles between two vectors, α

- and looking at distance between individual vectors (x, y) or data points in this vector space, ⇒ distance between xy

In order to measure angles and compute lengths and distances, we need to equip the vector space with an inner product, which will allow us to talk about geometric properties of elements in the vector space.

In next session we will look at:

- Computing lengths of vectors
- Computing distances between vectors
- Computing angles between vectors