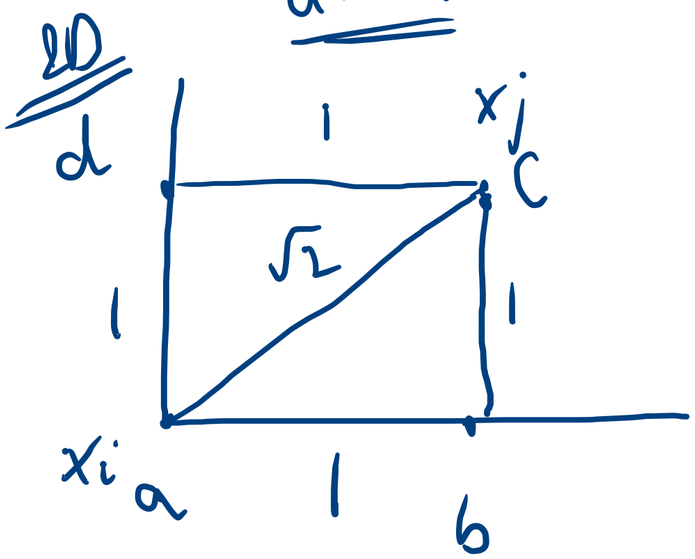
→ Class start at 9:05 pm

t-SNE (to preserve neigh. dist in d' dimension

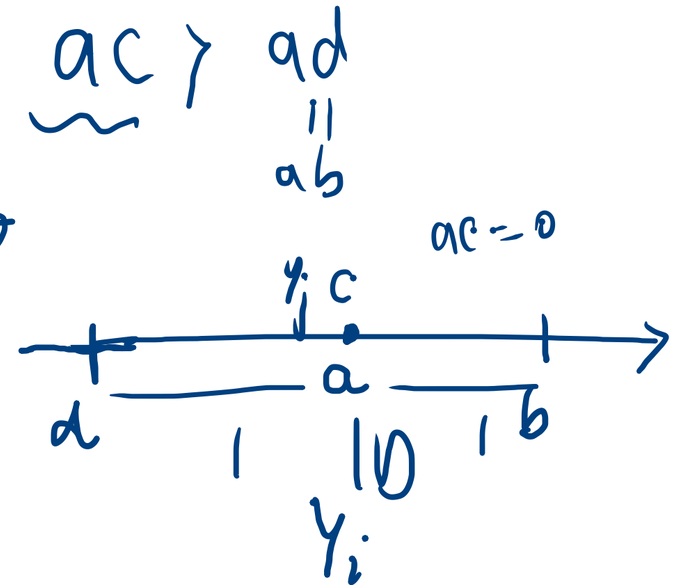
↳ t-distnuted Stochastic Neighbourhood Embedding

Crowding problem

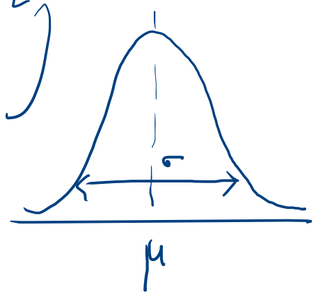
d-dim



d'-dim



$$p_i = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$



p_{ij} = dist b/w x_i & x_j

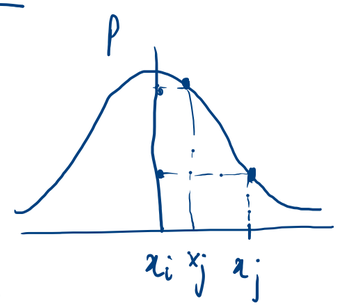
$p_{ij} \uparrow$ dist$(x_i, x_j) \downarrow$

$p_{ij} \downarrow$ dist$(x_i, x_j) \uparrow$

$$p_{ij} \propto \frac{1}{\text{dist}(x_i, x_j)}$$

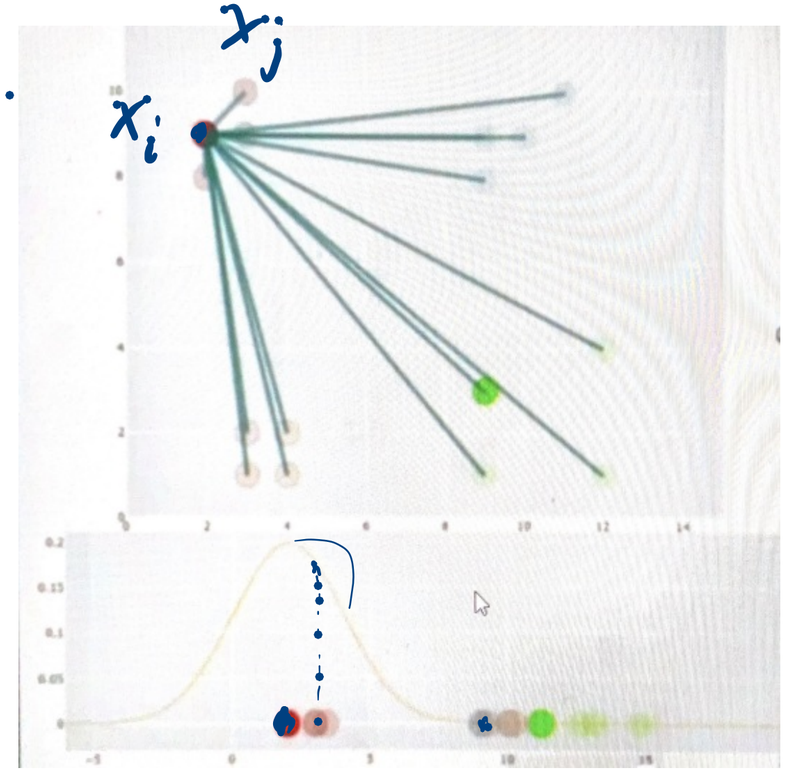
$$p_{ij} = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma_i^2}\right)$$

$$= \exp\left(\frac{-d_{ij}^2}{2\sigma_i^2}\right) \quad d_{ij} = \|x_i - x_j\|$$



$$p_{ij} = \frac{1}{\exp\left(\frac{d_{ij}^2}{2\sigma_i^2}\right)}$$

$k = 4$



$$p_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_i^2}\right)$$

$d_{ij} = f(\|x_i - x_j\|^2)$

= scalar quantity

In d-dimension

$$p_{i|j} = \frac{\exp\left(\|x_i - x_j\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq l} \exp\left(\frac{-\|x_k - x_l\|^2}{2\sigma^2}\right)}$$

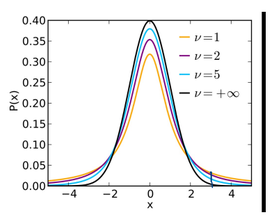
Pairwise similarity

$$\sum\sum p_{i|j} = 1$$

d-dimension → d'-dimension

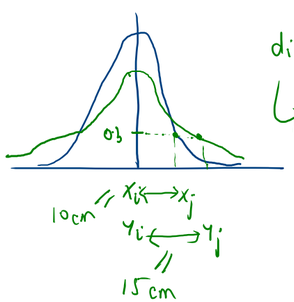
$p_{j|i}$ x_i, x_j → y_i, y_j $q_{j|i}$

Gaussian dist → t-distribution



→ t-dist has smaller peak and larger tail

→ t-dist with dof = 1



d_{ij} b/w x_i x_j

↳ $p_{j|i} = 0.3$ → d-dim

$q_{j|i} = 0.3$ → d'-dim

↳ d_{ij} b/w y_i & y_j is larger

→ Help somewhat overcome crowding problem.

$p_{j/i}$ → d-dim → gaussian

$q_{j/i}$ → d'-dim → t-dist with dof = 1

KL-divergence

$$KL(P,Q) = \sum_i \sum_j p_{ij} \log \left(\frac{p_{ij}}{q_{ij}} \right)$$

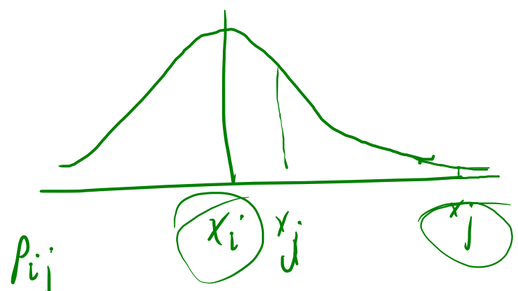
Compare 2 distribution and measure dissimilarity

$$\sum_{i=1}^{m} \sum_{j=1}^{c} p_{ij} \log \hat{p}_{ij}$$

① $p_{ij} = q_{ij}$ $\quad KL(P,Q) = 0$

② $KL \approx 0$ $\quad p_{ij} \approx 0$

x_i & x_j are far apart



p_{ij}

$p_{ij} \approx 0$

Perplexity → hyperparameter

↳ Effective no. of neigh who distance we want to preserve.

↳ 5 - 50

↳ Controlling local aspect and global aspect

Disadvantage

→ Very slow algorithm

→ Does not have predict function only

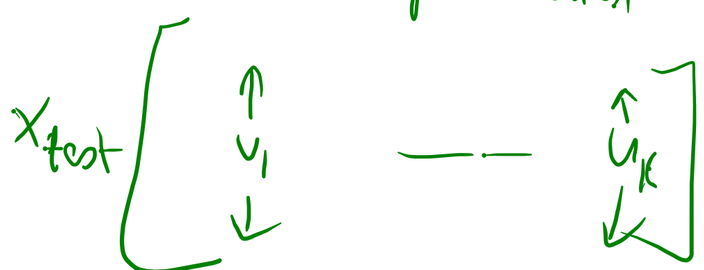
$D = \{x_i\}_{i=1}^{n} \longrightarrow D' = \{y_i\}_{i=1}^{n}$

x_{test} TSNE again y_{test}

p_{ij} q_{ij}

PCA

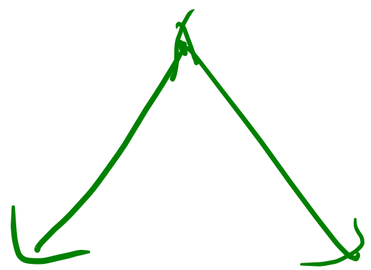
top-K eigen vectors



have top-K eigen values

→ Break until 10:25pm

UMAP → Uniform Manifold Approximation & Projection



Algebric topology (Manifold)

Algorithmic + optimization

→ Tries to preserve the structure as best as possible