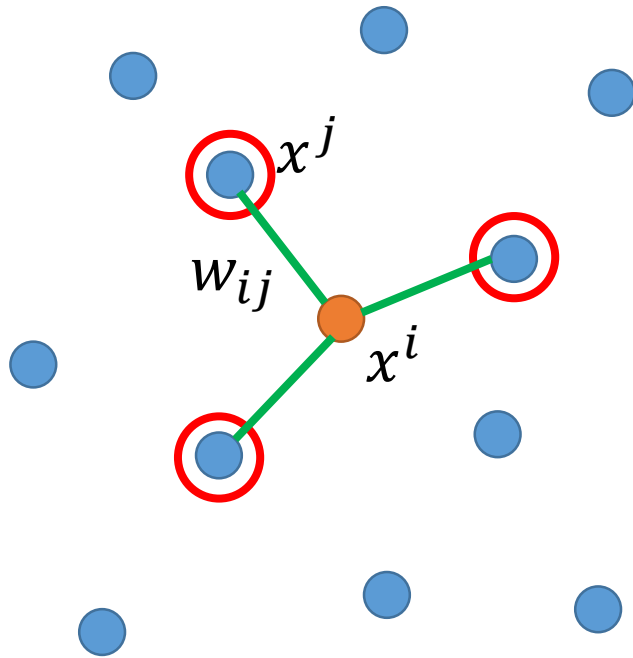


Unsupervised Learning: Neighbor Embedding

Locally Linear Embedding (LLE)



w_{ij} represents the relation between x^i and x^j

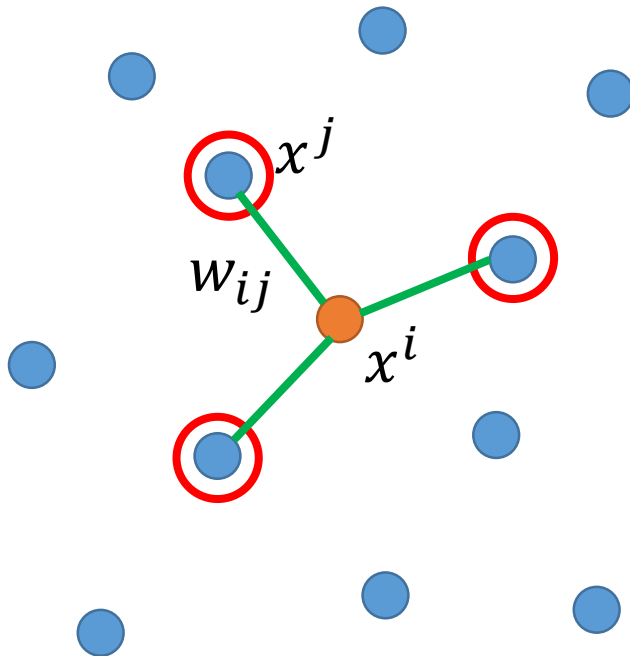
Find a set of w_{ij} minimizing

$$\sum_i \left\| x^i - \sum_j w_{ij} x^j \right\|_2$$

Then find the dimension reduction results z^i and z^j based on w_{ij}

LLE

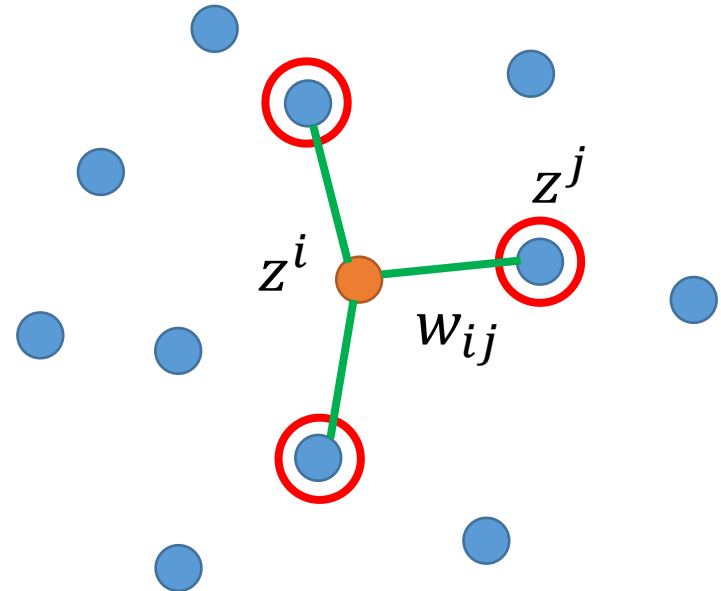
Keep w_{ij} unchanged



Original Space

Find a set of z^i minimizing

$$\sum_i \left\| z^i - \sum_j w_{ij} z^j \right\|_2$$

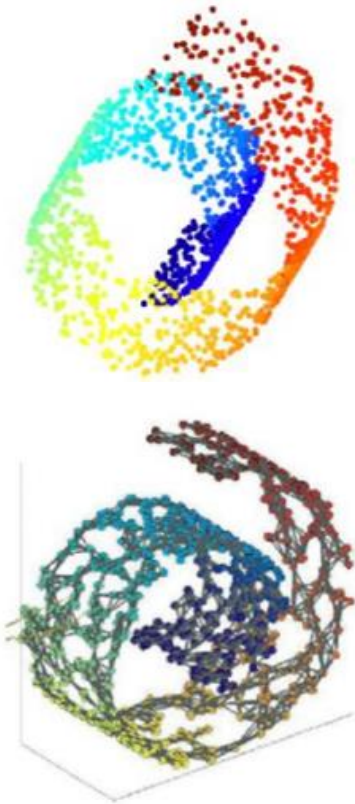


New (Low-dim) Space

Laplacian Eigenmaps

$$w_{i,j} = \begin{cases} \text{similarity} & \\ \text{If connected} & \\ 0 & \text{otherwise} \end{cases}$$

- *Review in semi-supervised learning:* If x^1 and x^2 are close in a high density region, \hat{y}^1 and \hat{y}^2 are probably the same.



$$L = \sum_{x^r} C(y^r, \hat{y}^r) + \lambda S$$

As a regularization term

$$S = \frac{1}{2} \sum_{i,j} w_{i,j} (y^i - y^j)^2 = \mathbf{y}^T L \mathbf{y}$$

S evaluates how smooth your label is

L: $(R+U) \times (R+U)$ matrix

Graph Laplacian

$$L = D - W$$

Laplacian Eigenmaps

- *Dimension Reduction*: If x^1 and x^2 are close in a high density region, z^1 and z^2 are close to each other.

$$S = \frac{1}{2} \sum_{i,j} w_{i,j} (z^i - z^j)^2$$

Any problem? How about $z^i = z^j = \mathbf{0}$?

Giving some constraints to z :

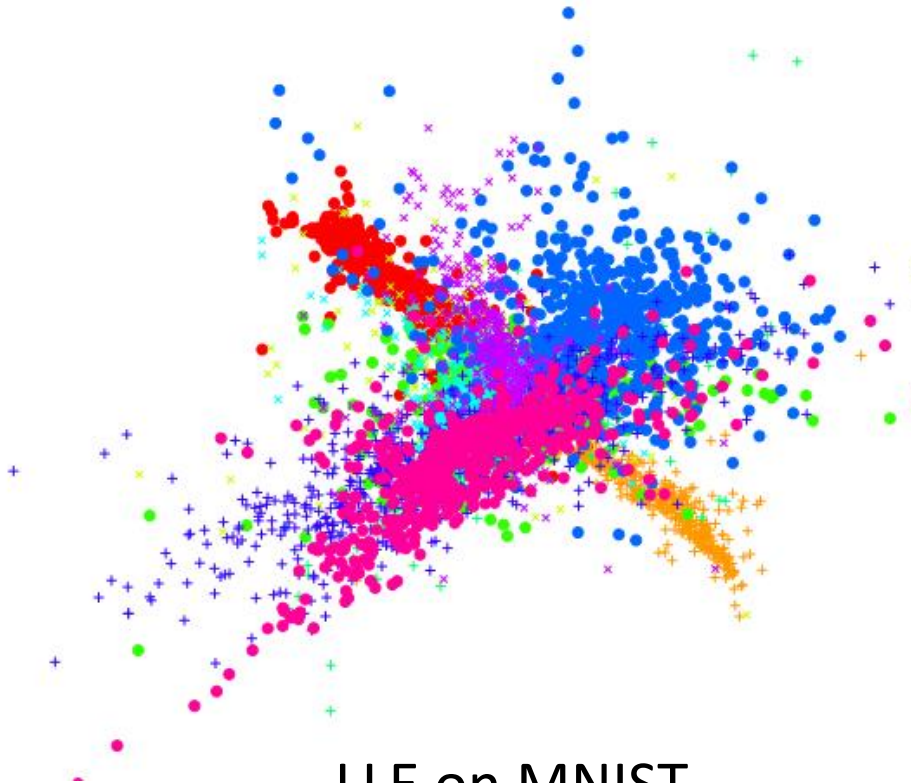
If the dim of z is M , $\text{Span}\{z^1, z^2, \dots, z^N\} = \mathbb{R}^M$

Spectral clustering: clustering on z

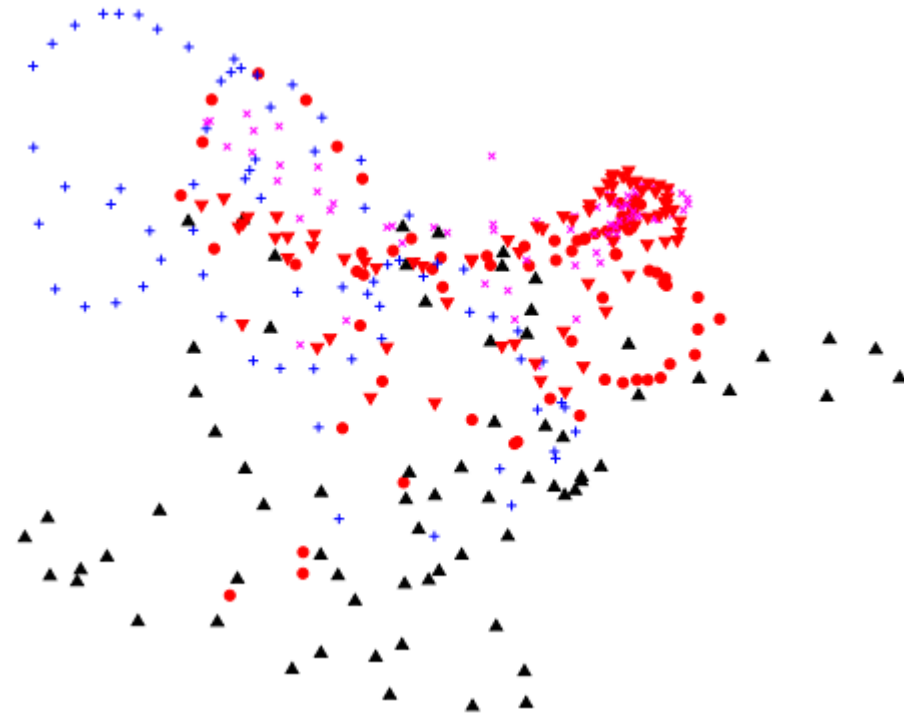
Belkin, M., Niyogi, P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Advances in neural information processing systems* . 2002

T-distributed Stochastic Neighbor Embedding (t-SNE)

- Problem of the previous approaches
 - Similar data are close, but different data may collapse



LLE on MNIST



LLE on COIL-20

t-SNE



Compute similarity between all pairs of x : $S(x^i, x^j)$

$$P(x^j | x^i) = \frac{S(x^i, x^j)}{\sum_{k \neq i} S(x^i, x^k)}$$

Compute similarity between all pairs of z : $S'(z^i, z^j)$

$$Q(z^j | z^i) = \frac{S'(z^i, z^j)}{\sum_{k \neq i} S'(z^i, z^k)}$$

Find a set of z making the two distributions as close as possible

$$\begin{aligned} L &= \sum_i KL(P(* | x^i) || Q(* | z^i)) \\ &= \sum_i \sum_j P(x^j | x^i) \log \frac{P(x^j | x^i)}{Q(z^j | z^i)} \end{aligned}$$

Ignore σ for
simplicity

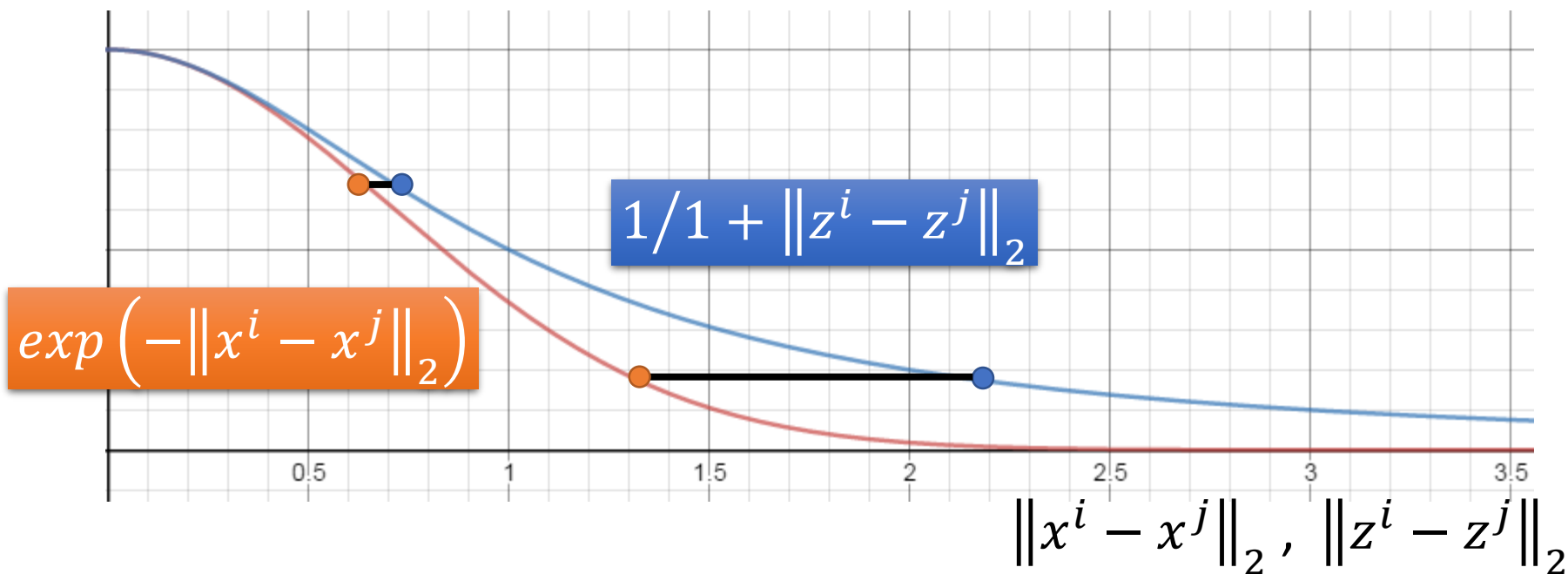
t-SNE – Similarity Measure

SNE:

$$S(x^i, x^j) = \exp(-\|x^i - x^j\|_2)$$

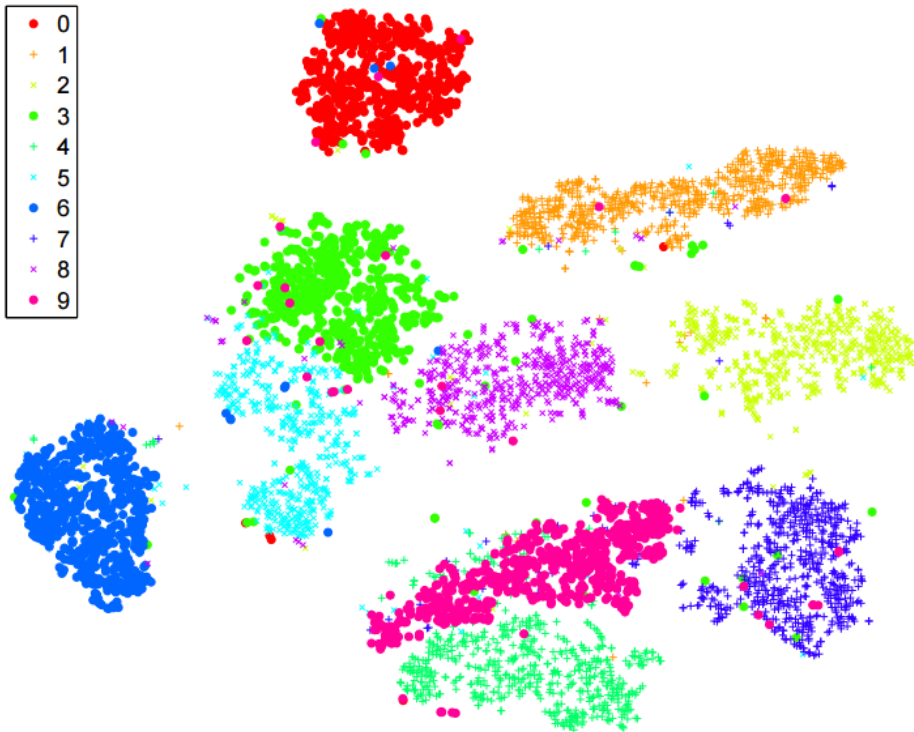
t-SNE:

$$S'(z^i, z^j) = 1 / (1 + \|z^i - z^j\|_2)$$



t-SNE

- Good at visualization



t-SNE on MNIST



t-SNE on COIL-20