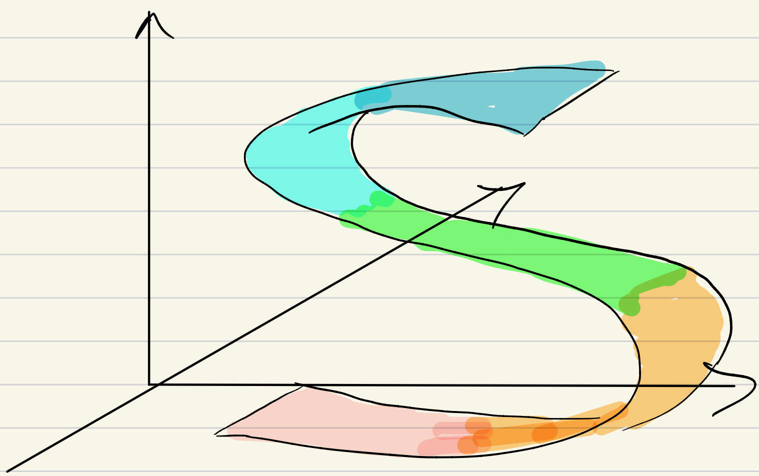
Lecture 16

- Non-linear Dimensional Reduction
  ("Manifold Learning")

⇛ PCA, NMF, and ICA are great at first trying dimensional reduction

⇛ BUT there's no reason why the physics generating our data (and thus our data features) should be LINEAR ... MANY PROBLEMS ARE NON-LINEAR!



* Imagine we have data drawn from this 2D S-shaped manifold embedded in 3D.
  ↳ colors denote target labels

* PCA will NOT be able to linearly separate.

⇛ BUT manifold-learning techniques can learn the geometry of the data, and UNWRAP this.

- Locally Linear Embedding (LLE)

  * Embed high-D space in lower-D space.

  * Preserve geometry of local "neighborhoods"

1. Define local geometry

   don't allow $W = I$

   ↳ minimize $\mathcal{E}_1(W) = |X - WX|^2$ to find W.

   ↳ "W" is weights matrix that constructs a point from "k" neighbors.

$X = N \times K$ matrix

2. Embed in lower-D space.

   ↳ $\mathcal{E}_2(Y) = |Y - WY|^2$

   ↳ Y has lower dimension than X, but weights matrix W is fixed.

$Y = N \times d$ matrix ($d \ll K$)

Isometric Mapping (IsoMap)

* IsoMap actually computes distances between points along geodesics of the manifold's geometry.
  not Euclidean distances.

* IsoMap then finds a lower-D embedding that preserves the geodesic distances between points ⇒ structure is maintained

t-distributed Stochastic Neighbor Embedding (t-SNE)

* t-SNE preserves only local similarities, not global properties (which PCA tries).

1. Measure the similarity between nearby points in high-dimensional space.
   ↳ evaluate the probability of each point under a multi-variate Gaussian centered on each point... p_{ij}

2. Find a lower-D mapping that preserves p_{ij}
   ↳ dissimilar points are far apart on the lower-D embedding.