Chapter18 Unsupervised Learning – Neighbor Embedding

1. Manifold Learning
2. Locally Linear Embedding (LLE)
   a) w_{ij} represents the relation between x^i and x^j
   b) Find a set of w_{ij} minimizing $\sum_i ||x^i - \sum_j w_{ij}x^j||_2$
   c) Then, find the dimension reduction results z^i and z^j based on w_{ij}
   d) 在天愿做比翼鸟，在地愿为连理枝
   e) Keep w_{ij} unchanged, find a set of z^i minimizing $\sum_i ||z^i - \sum_j w_{ij}z^j||_2$
3. Laplacian Eigenmaps
   a) Graph-based approach: construct the data points as a graph
   b) If x^1 and x^2 are close in a high density region, z^1 and z^2 close to each other
   c) $S = \frac{1}{2}\sum_{i,j} w_{ij}(z^i - z^j)^2$ Any problem? How about $z^i = z^j = 0$?
   d) Given some constraints to z:
      If the dim of z is M, $Span\{z^1, z^2, \dots, z^N\} = R^M$
4. T-distributed Stochastic Neighbor Embedding (t-SNE)
   a) Problem of the previous approaches:
      Similar data are close, but different data may collapse
   b) Compute similarity between all pairs of x and z
      $P(x^j|x^i) = \frac{S(x^i,x^j)}{\sum_{k!=i} S(x^i,x^k)}$ $Q(z^j|z^i) = \frac{S\prime(z^i,z^j)}{\sum_{k!=i} S\prime(z^i,z^k)}$
   c) Find a set of z making the two distributions as close as possible

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

   d) SNE: $S(x^i, x^j) = \exp(-||x^i - x^j||_2)$
      t-SNE: $S'(z^i, z^j) = \frac{1}{1+||z^i - z^j||_2}$