# w2 tSNE

Ngọc Anh Phạm

January 2022

## 1 Xây dựng lại bài toán t-SNE. Tính đạo hàm loss với các parameter (y) trong bài toán t-SNE

SNE converts euclidean distances to similarities, that can be interpreted as probabilities

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

$$q_{j|i} = \frac{exp(-||y_i - y_j||^2)}{\sum_{k\neq i} exp(-||y_i - y_k||^2)}$$

$$p_{j|i} = 0, q_{j|i} = 0$$

t-SNE minimizes the Kullback-Leibler divergence between the joint probabilities $p_{ij}$ in the high-dimensional space and the joint probabilities $q_{ij}$ in the low-dimensional space. The values of $p_{ij}$ are defined to be the symmetrized conditional probabilities, whereas the values of $q_{ij}$ are obtained by means of a Student-t distribution with one degree of freedom:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}$$

$$q_{ij} = \frac{(1 + ||y_i - y_j||^2)^{-1}}{\sum_{k\neq l}(1 + ||y_k - y_l||^2)^{-1}}$$

The Kullback-Leibler divergence between the two joint probability distributions P and Q is given by:

$$C = \sum_i KL(P_i||Q_i) = \sum_i \sum_j p_{j|i} log \frac{p_{j|i}}{q_{j|i}}$$

$$= \sum_i \sum_j p_{ij} \log p_{ij} - p_{ij} \log q_{ij}$$

In order to make the derivation less cluttered, we define two auxiliary variables $d_{ij}$ and Z as follows:

$$d_{ij} = ||y_i - y_j||$$

$$Z = \sum_{k \neq l} = (1 + d_{kl}^2)^{-1}$$

Note that if $y_i$ changes, the only pairwise distance that change are $d_{ij}$ and $d_{ji}$ for $\forall j$. Hence, the gradient of the cost function C with respect to $y_i$ is given by

$$\frac{\delta \mathrm{C}}{\delta y_i} = \sum_j (\frac{\delta \mathrm{C}}{\delta d_{ij}} + \frac{\delta \mathrm{C}}{\delta d_{ij}})(y_i - y_j)$$

$$= 2 \sum_j \frac{\delta \mathrm{C}}{\delta d_{ij}} (y_i - y_j) \qquad (1)$$

The gradient $\frac{\delta \mathrm{C}}{\delta d_{ij}}$ The gradient $\frac{\delta C}{\delta d_{ij}}$ is computed from the definition of the Kullback-Leibler divergence:

$$\frac{\delta C}{\delta d_{ij}} = -\sum_{k \neq l} p_{kl} \frac{\delta (log q_{kl})}{\delta d_{ij}}$$

$$= -\sum_{k \neq l} p_{kl} \frac{\delta (log q_{kl} Z - log Z)}{\delta d_{ij}}$$

$$= -\sum_{k \neq l} p_{kl} (\frac{1}{q_{kl} Z} \frac{\delta ((1 + d_{kl}^2)^{-1})}{\delta d_{ij}} - \frac{1}{Z} \frac{\delta Z}{\delta d_{ij}})$$

The gradient $\frac{\delta ((1+d_{kl}^2)^{-1})}{\delta d_{ij}}$ is only nonzero when k = i and l = j. Hence, the gradient $\frac{\delta C}{\delta d_{ij}}$ is given by

$$\frac{\delta C}{\delta d_{ij}} = 2 \frac{p_{ij}}{p_{ij} Z} (1 + {d_{ij}}^2) - 2 \sum_{k \neq 1} p_{kl} \frac{(1 + {d_{ij}}^2)^{-2}}{Z}$$

Noting that $\sum_{k \neq l} p_{kl} = 1$, we see that the gradient simplifies to

$$\frac{\delta C}{\delta d_{ij}} = 2 p_{ij} (1 + d_{ij}^2)^{-1} - 2 q_{ij} (1 + d_{ij}^2)^{-1}$$

$$= 2 (p_{ij} - q_{ij}) (1 + d_{ij}^2)^{-1}$$

Substituting this term into (1), we obtain the gradient:

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij}) (1 + ||y_i - y_j||^2)^{-1} (y_i - y_j)$$