Stochastic Neighbor Embedding (SNE)

Pham Van Hong

11201638

Problem 1:

Definition: (SNE)

$$q_{j|i} = \frac{e^{-\left\|y_i - y_j\right\|^2}}{e^{-\left\|y_i - y_j\right\|^2}} = \frac{E_{ij}}{E_{ik}} = \frac{E_{ij}}{Zi}$$

And $E_{ij} = E_{ji}. Conclusion, the\ loss\ function\ is\ defined\ as$:

$$C = \sum_{k,l \neq k} p_{l|k} \log \frac{p_{l|k}}{q_{l|k}} = \sum_{k,l \neq k} p_{(l|k)} \log p_{l|k} - p_{l|k} \log q_{l|k}$$

$$C = \sum_{k,l \neq k} p_{l|k} \log p_{l|k} - p_{l|k} \log E_{lk} + p_{l|k} \log Z_k$$

We derive with respect to y_i:

$$\frac{\partial C}{\partial y_i} = \sum_{k,l \neq k} -p_{l|k} \partial \log E_{kl} + \sum_{k,l \neq k} p_{l|k} \partial \log Z_k$$

$$\sum_{k,l \neq k} - p_{l|k} \partial \log E_{k|l} = \sum_{j \neq i} -p_{j|i} \partial \log E_{ij} - p_{i|j} \partial \log E_{ji}$$

Since $E_{ij} = E_{ij} \left(-2\left(y_i - y_j\right)\right) we\ have$

$$\sum_{j \neq i} -\frac{p_{j|i} E_{ij}}{E_{ij}} \left(-2\left(y_i - y_j\right)\right) - \frac{p_{i|j} E_{ji}}{E_{ji}} \left(2\left(y_j - y_i\right)\right)$$

$$= 2 \sum_{j \neq i} \left(p_{j|i} + p_{i|j}\right)(y_i - y_j)$$

Since $\sum_{l \neq j} p_{l|j} = 1\ and\ Z_j\ does\ not\ depend\ on\ k, we\ can\ write$

$$\sum_{j,k\neq j} p_{k|j}\,\partial \log Z_j = \sum_j \partial\log Z_j$$

$$= \sum_j \frac{1}{Z_j}\sum_{k\neq j}\partial E_{jk}$$

$$= \sum_j \frac{E_{ji}}{Z_j}\Big(2\big(y_j - y_i\big)\Big) + \sum_{j\neq i}\frac{E_{ij}}{Z_i}\Big(-2\big(y_i - y_j\big)\Big)$$

$$= 2\sum_{j\neq i}\big(-q_{j|i} - q_{i|j}\big)\big(y_i - y_j\big)$$

$$C = 2\sum_{j\neq i}(p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j})(y_i - y_j)$$

t-distributed SNE

Definition:

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

$$= \frac{E_{ij}^{-1}}{\sum_{k,l\neq k} E_{kl}^{-1}}$$

$$= \frac{E_{ij}^{-1}}{Z}$$

We have $E_{ij} = E_{ji}, the\ loss\ function\ is\ defined\ as$

$$C = \sum_{k,l\neq k} p_{lk}\log\frac{p_{lk}}{q_{lk}} = \sum_{k,l\neq k} p_{lk}\log p_{lk} - p_{lk}\log q_{lk}$$

$$= \sum_{k,l\neq k} p_{lk}\log p_{lk} - p_{lk}\log E_{kl}^{-1} + p_{lk}\log Z$$

We derive with respect to y_i

$$\frac{\partial C}{\partial y_i} = \sum_{k,l\neq k} -p_{lk}\partial\log E_{kl}^{-1} + \sum_{k,l\neq k} p_{lk}\partial\log Z$$

We start with the first term, noting that the derivative is non-zero when ∀j, k = i or l = i, that $p_{ij} = p_{ji}$and $E_{ij} = E_{ji}$

$$\sum_{k,l \neq k} -p_{lk} \partial \log E_{kl}^{-1} = -2 \sum_{j \neq i} p_{ij} \partial \log E_{ij}^{-1}$$

Since $\partial E_{ij}^{-1} = E_{ij}^{-2}\left(-2(y_i - y_j)\right) we\ have$:

$$-2 \sum_{j \neq i} \frac{p_{ji} E_{ij}^{-2}}{E_{ij}^{-1}} \left(-2(y_i - y_j)\right) = 4 \sum_{j \neq i} p_{ji} E_{ij}^{-1} (y_i - y_j)$$

$$\sum_{k,l \neq k} p_{lk} \partial \log Z = \frac{1}{Z} \sum_{k',l' \neq k'} \partial E_{kl}^{-1}$$

$$= 2 \sum_{k,l \neq k} p_{lk} \partial \log Z$$

$$= \frac{1}{Z} \sum_{k',l' \neq k'} \partial E_{kl}^{-1}$$

$$= 2 \sum_{j \neq i} \frac{E_{ji}^{-2}}{Z} \left(-2(y_j - y_i)\right)$$

$$= \ -4 \sum_{j \neq i} q_{ij} E_{ji}^{-1} (y_i - y_j)$$

$$= \frac{\partial C}{\partial y_i} = 4 \sum_{j \neq i} (p_{ji} - q_{ji}) E_{ji}^{-1} (y_i - y_j)$$

$$= \frac{\partial C}{\partial y_i} = 4 \sum_{j \neq i} (p_{ji} - q_{ji}) \left(1 + \left|\left|y_i - y_j\right|\right|^2\right)^{-1} (y_i - y_j)$$