Nguyễn Quốc Huy - 11205484.

we derive with respect to y_i. To make the derivation less cluttered, I will omit the ∂y_i term at the denominator

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

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

$$\sum_{k,l\neq k} -P_{lk}\,\partial \log E_{kl}^{-1} = -2\sum_{j\neq i} P_{ji}\,\partial \log E_{ij}^{-1} \qquad (9)$$

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

$$-2\sum_{j\neq i} P_{ji}\frac{E_{ij}^{-2}}{E_{ij}^{-1}}(-2(y_i - y_j)) = 4\sum_{j\neq i} P_{ji}E_{ij}^{-1}(y_i - y_j) \qquad (10)$$

we conclude with the second term. Using the fact that $\sum_{k,l\neq k} P_{kl} = 1$ and that Z does not depend on k or l

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

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

$$= -4\sum_{j\neq i} q_{ij}E_{ji}^{-1}(y_i - y_j) \qquad (11)$$

Combining (10) and (11)

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

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

Nguyen Quoc Huy - 11205484

we conclude with second term, since $\sum_{l \neq j} P_{l|j} = 1$ and Z_j does not depend on k, we can write

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

The derivative is non-zero when $k = i$ or $j = i$

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

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

$$= 2 \sum_{j \neq i} (-q_{j|i} - q_{i|j})(y_i - y_j) \quad (5)$$

Combining (4) and (5) we have

$$\frac{\partial C}{\partial y_i} = 2 \sum_{j \neq i} (p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j})(y_i - y_j) \quad (6)$$

$\boxed{\text{t - SNE}}$

Define

$$q_{ji} = q_{ij} = \frac{(1 + \| y_i - y_j \|^2)^{-1}}{\sum_{k, l \neq k} (1 + \| y_k - y_l \|^2)^{-1}} = \frac{E_{ij}^{-1}}{\sum_{k, l \neq k} E_{kl}^{-1}} = \frac{E_{ij}^{-1}}{Z} \quad (7)$$

Notice that $E_{ij} = E_{ji}$. The loss function:

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

$$= \sum_{k, l \neq k} P_{lk} \log P_{lk} - P_{lk} \log E_{kl}^{-1} + P_{lk} \log Z \quad (8)$$

Scanned with CamScanner

Nguyễn Quốc Huy - 11205484.

1

Question 1: $\boxed{\text{SNE}}$

Define

$$q_{j|i} = \frac{e^{-\|y_i - y_j\|^2}}{\sum_{k \neq i} e^{-\|y_i - y_k\|^2}} = \frac{E_{ij}}{\sum_{k\neq i} E_{ik}} = \frac{E_{ij}}{Z_i} \quad (1)$$

Notice that: $E_{ij} = E_{ji}$. The loss function is defined as

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

$$= \sum_{k, l\neq k} P_{l|k} \log P_{l|k} - p_{l|k} \log E_{kl} + P_{l|k} \log Z_k \quad (2)$$

We derive with respect to y_i. To make the derivation less cluttered, I will omit the ∂y_i term at the denominator.

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

we start with the first item, ~~noticing~~ note that the derivative is non-zero when $\forall j \neq i$, $k = i$ or $l = i$

$$\sum_{k, l\neq k} -P_{l|k} \partial \log E_{kl} = \sum -P_{j|i} \partial \log E_{ij} - P_{i|j} \partial \log E_{ji} \quad (3)$$

since $\partial E_{ij} = E_{ij}(-2(y_i - y_j))$ we have

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

$$= 2 \sum_{j\neq i} (P_{j|i} + P_{i|j})(y_i - y_j) \quad (4)$$