Homework week 2

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

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

with $p_{i|i} = q_{i|i} = 0$
Loss calculation
Define:

$$p_{i|j} = q_{j|i} = \frac{(1 + \|y_i - y_j\|)^{-1}}{\Sigma_{k,l\neq k}(1 + \|y_i - y_k\|)^{-1}} = \frac{E_{ij}^{-1}}{\Sigma_{k,l\neq k} E_{kl}^{-1}} = \frac{E_{ij}^{-1}}{Z}$$

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

$$C = KL(P\|Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ji}}{q_{ji}}$$

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

We drive with respect to y_i

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

The derivative is non-zero when j, $k = i$ or $l = j$, that $p_{ji} = p_{ij}$ and $E_{ji} = E_{ij}$

$$\sum_{k,l\neq k} -p_{lk} \frac{\delta C \log E_{kl}^{-1}}{\delta C y_i} = -2 \sum_{j\neq i} -p_{ji} \frac{\delta C \log E_{ij}^{-1}}{\delta C y_i}$$

Since $\frac{\delta C \log E_{ij}^{-1}}{\delta C y_i} = E_{ij}^{-2}(-2(y_i - y_j))$, then

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

Using $\sum_{k,l\neq k} p_{kl} = 1$ and Z doesn't depend on k or l

$$\sum_{k,l\neq k} p_{kl}\frac{\delta \log Z}{\delta y_i} = \frac{1}{Z}\sum_{k',l'\neq k'}\frac{\delta E_{kl}^{-1}}{\delta y_i} = 2\sum_{j\neq i}\frac{\delta E_{ji}^{-2}}{Z}(-2(y_j - y_i)) = -4\sum_{j\neq i} q_{ij}E_{ji}^{-1}(y_i - y_j)$$

Final result:

$$\frac{\delta C}{\delta y_i} = 4\sum_j \left(p_{ji} - q_{ji}\right) E_{ji}^{-1} \left(y_i - y_j\right)$$

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