ML2 WEEK2-latex

nguyen thi phuong

January 2022

1 Introduction

$X = x_1, x_2, ...x_n$ in R^h in high dimension $->Y y_1, y_2...y_n$ in R^l with $l < h$

we want to convert eculidean distaces to similarities that can be interpreted as probabilities with:

$p_j/_i = \frac{\exp(\frac{-||x_i - x_j||^2}{\sigma^2})}{\sum_k \exp(\frac{-||x_i - x_k||^2}{\sigma^2})}$ and $q_j/_i = \frac{\exp(-||y_i - y_j||^2)}{\sum_k \exp(-||y_i - y_k||^2)}$ with $p_i/_i - 0, q_i/_i = 0$
$KL(P_i||Q_i) = \sum_i \sum_j (p_j/i \log(\frac{p_j/_i}{q_j/_i}))$

$KL(P_i||Q_i) = \sum_i \sum_j (p_j/i. \log(p_j/_i) - p_j/_i \log(q_j/_i))$

Let $\sum_i \sum_j (p_j/i. \log(p_j/_i)) = A$

$KL(P_i||Q_i) = \sum_i \sum_j A - p_j/_i \log(q_j/_i))$

Now, we will take the derivative of $\sum_i \sum_j -p_j/_i \log(q_j/_i)$

$$D = \sum_i \sum_j -p_j/_i \log(\frac{\exp(-||y_i - y_j||^2)}{\sum_k \exp(-||y_i - y_k||^2))}$$

For each i, we have to take derivative of for example of i=1 then $q_1/_1, q_2/_1, ...q_n/_1$ and $q_1/_1, q_1/_2, ...q_1/_n$

$D = \sum_i \sum_j -p_j/_i. \log(q_j/_i) - p_i/_j.log(q_i/_j)$

We will temporarily let $p_j/_i$,$p_i/_j$ out of each side of the above term and take derivative of the remain. We take derivative side by side. For each side, Ex:

$\sum_i \sum_j - \log(q_j/_i) = \sum_i \sum_j - \log(\exp(-||y_i - y_j||^2) + \log(\sum_k \exp(-||y_i - y_k||^2))$

Then, for each i, we get:

$$\frac{d(D)}{d(y_i)} = \sum_j (2(y_i - y_j) + \frac{-2\sum_j (y_i - y_j)\exp(-||y_i - y_j||^2)}{\sum_j \exp(-||y_i - y_j||^2)} + \sum_j (2(y_j - y_i) + \frac{-2\sum_j (y_j - y_i)\exp(-||y_j - y_i||^2)}{\sum_j \exp(-||y_j - y_i||^2)}$$

$$\frac{d(D)}{d(y_i)} = \sum_j (2(y_i - y_j) - 2\sum_j (y_i - y_j).q_j/_i + \sum_j (2(y_j - y_i) - 2\sum_j (y_j - y_i).q_i/_j$$

Now, we multiple $p_i/_j$ and $p_j/_i$ into the above equation, we get:

$$\frac{d(D)}{d(y_i)} = \sum_j (2(y_i - y_j).p_j/_i - 2\sum_j (y_i - y_j).q_j/_i.p_j/_i + \sum_j (2(y_j - y_i).p_i/_j - 2\sum_j (y_j - y_i).q_i/_j.p_i/_j$$