t-SNE Notes

Pavlin Poličar

1 KL Divergence for exaggerated P

$$KL(P \mid\mid Q) = \sum_{ij} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{1}$$

We need to introduce the scaling i.e. exaggeration factor α

$$= \sum_{ij} \frac{\alpha}{\alpha} p_{ij} \log \frac{\alpha p_{ij}}{\alpha q_{ij}} \tag{2}$$

Exaggeration means that the p_{ij} terms get multiplied by α, so we need to find an expression for the KL divergence that includes only αp_{ij} and q_{ij} and some other factor that will correct for α.

$$= \frac{1}{\alpha} \sum_{ij} \alpha p_{ij} \left(\log \frac{\alpha p_{ij}}{q_{ij}} - \log \alpha \right) \tag{3}$$

$$= \frac{1}{\alpha} \left(\sum_{ij} \alpha p_{ij} \log \frac{\alpha p_{ij}}{q_{ij}} - \sum_{ij} \alpha p_{ij} \log \alpha \right) \tag{4}$$

The first term is computed by the negative gradient methods (since they don't know about exaggeration, and only care about P, the second term can easily be