Kullback - Leibler Divergence

(KLD)

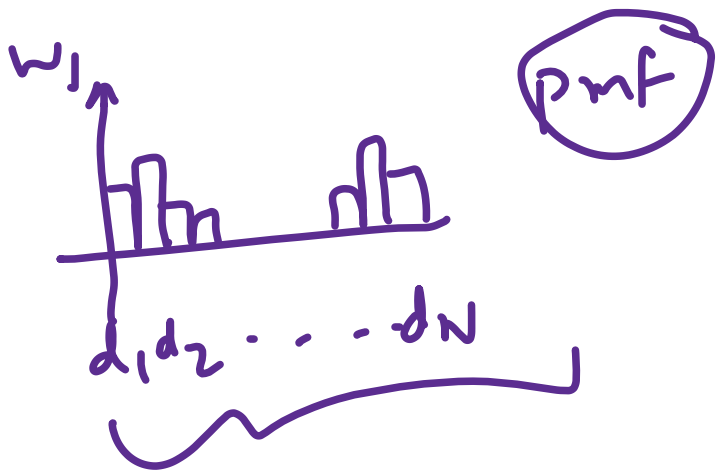
distance between distributions

$$KLD(P,Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

Relative entropy =

$$\sum_x P(x) \log P(x) - \sum_x P(x) \log Q(x)$$

formalism



$$KLD(P,Q) = D(P\|Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

$$= -\sum_x P(x) \log Q(x) - \left(-\sum_x P(x) \log P(x)\right)$$

optimal

Entropy

Suboptimal

Relative Entropy

Q is a surrogate of P

≥ 0

$P(x) \quad \log \frac{1}{P(x_i)}$

$KLD = 0 \quad \text{iff } P = Q$

KLD
1) asymmetric
2) $KLD(P,Q) = 0$ iff $P = Q$
3) $KLD(P,Q) \geq 0$

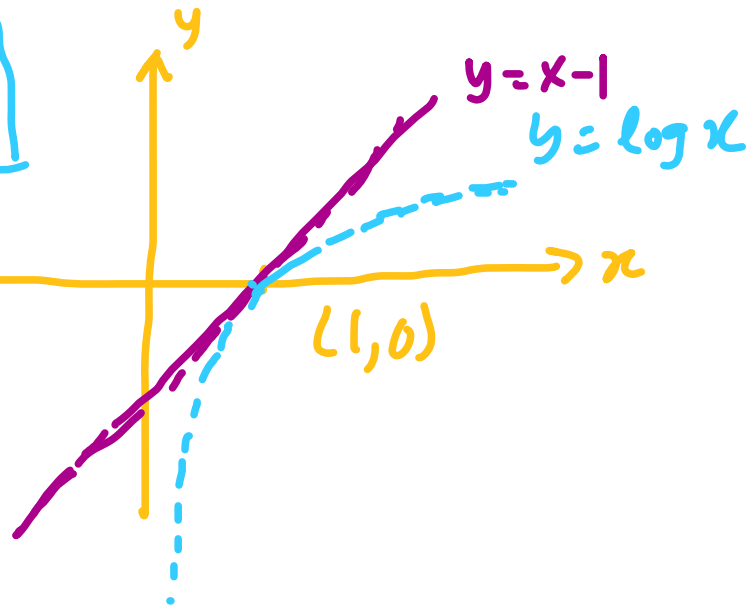
KLD ≥ 0

$$-KLD(P,Q) = -\sum_{x} P(x) \log \frac{P(x)}{Q(x)} = \sum_{x|P(x)>0} P(x) \log \frac{Q(x)}{P(x)}$$

$$\leq \sum_{x|P(x)>0} P(x) \left[\frac{Q(x)}{P(x)} - 1 \right]$$

Jensen's Inequality

$$\log x \leq x - 1$$



$$= \underbrace{\sum_{x|P(x)>0} Q(x)}_{\leq 1} - \underbrace{\sum_{x|P(x)>0} P(x)}_{1}$$

$$\leq 0$$

KLD(P,Q) ≥ 0 $\Leftarrow$

Symmetric

Jensen-Shannon Divergence (JSD)

$$\frac{1}{2}\left[D(P\|M) + D(Q\|M)\right]$$

$$M = \frac{P+Q}{2}$$

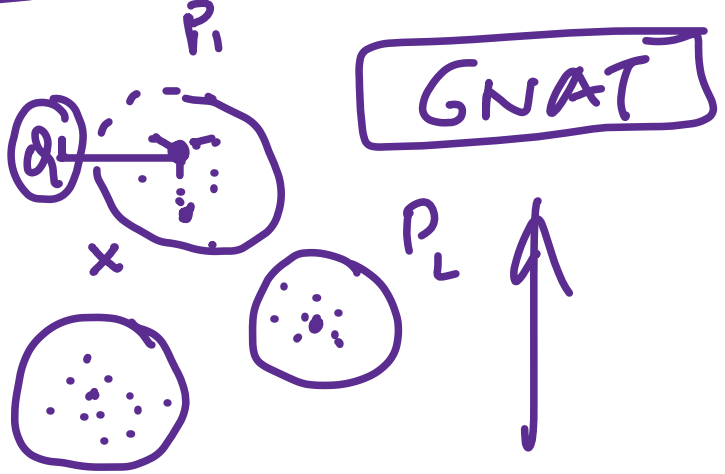
$\sqrt{JSD}$

- metric → non-negativity, identity, Symmetry, triangle inequality.

$$d(Q,X) \leq d(Q,P) + d(P,X)$$

$$d(Q,X) \geq |d(Q,P) - d(P,X)|$$

efficiency



GNAT

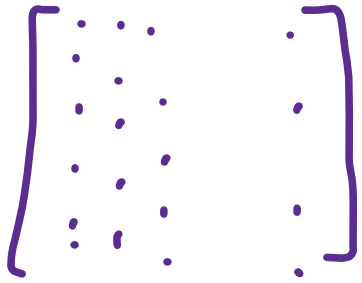
P3 Geometric Near
Neighbour Access Tree

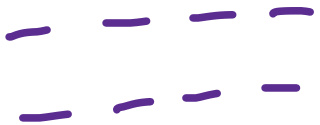
GRADE UPGRADE QUESTION

→ LSA
PLSA
KLD
Word Embeddings.

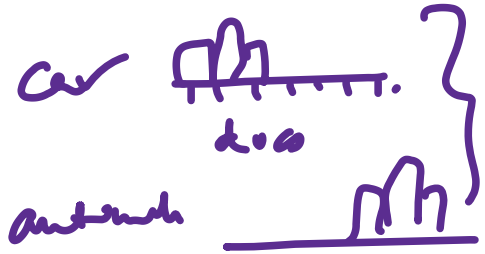
Why is anything more sophisticated needed at all?

term term

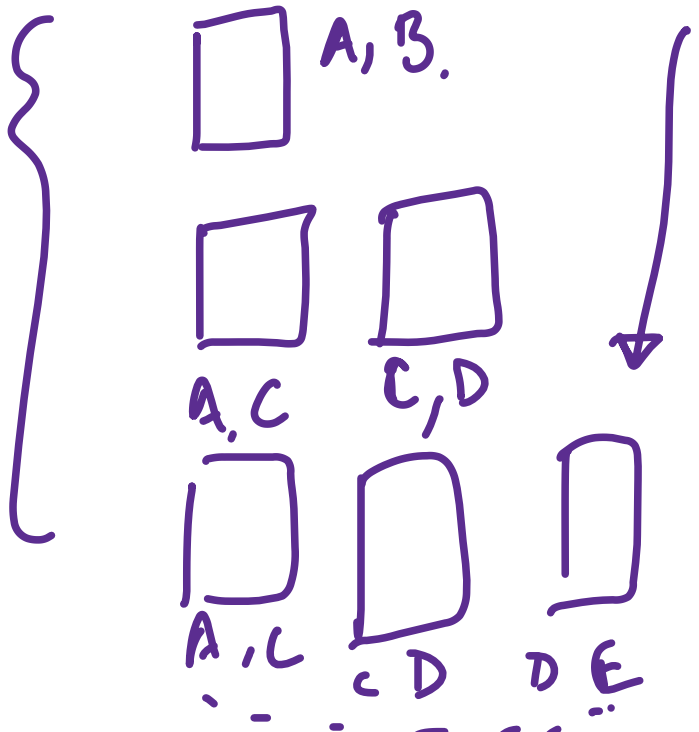




KLD car dog

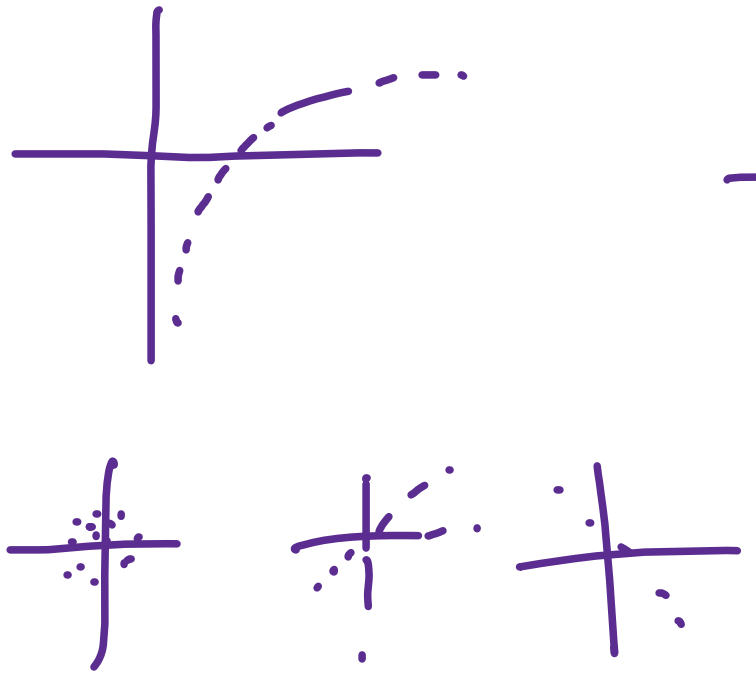


Orders of association.



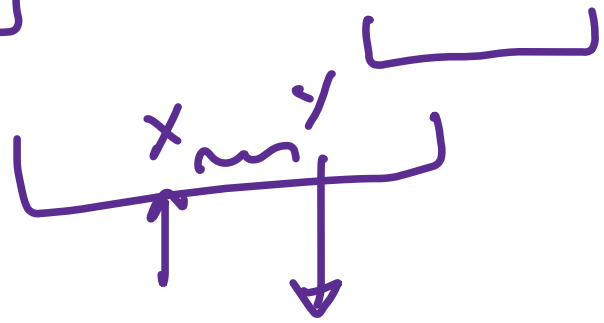
Intuition of negative PMI

$$\log \left[\frac{P(x,y)}{P(x) \cdot P(y)} \right]$$



$$\left[\log \frac{1}{P(y)} \right]$$

$$P(x|y) = P(x)$$



⊗