c∗GMVÆs—q

Kolb, Yiftach

January 19, 2023

MAX-PLANCK-INSTITUT
FÜR MOLEKULARE GENETIK

c∗GMVÆs—q

a Master Thesis in Bioinformatics

Advisor / Reviewer: Professor Martin Vingron
Reviewer: Professor Tim Conrad

Freie Universität Berlin

Autoencoders

Why turbolence

VAEs

Why turbolence

$$
\begin{aligned}
\frac{1}{N}\log p(\boldsymbol{X}) &= \frac{1}{N}\log\int p(\boldsymbol{X},\boldsymbol{Z})d\boldsymbol{Z} && \text{taking marginal}\\
&= \frac{1}{N}\log\int \frac{p(\boldsymbol{X},\boldsymbol{Z})}{q(\boldsymbol{Z})}q(\boldsymbol{Z})d\boldsymbol{Z} && \text{multiplying by 1 inside}\\
&= \frac{1}{N}\log\int \frac{p(\boldsymbol{X},\boldsymbol{Z})}{q(\boldsymbol{Z})}dq(\boldsymbol{Z}) && \text{definition of } dq(\mathbf{Z})\\
&\geq \frac{1}{N}\int \log\frac{p(\boldsymbol{X},\boldsymbol{Z})}{q(\boldsymbol{Z})}dq(\boldsymbol{Z}) && \text{Jensen inequality}\\
&= \frac{1}{N}\int\sum_{1}^{N}\log\frac{p(\mathbf{x}_i,\mathbf{z}_i)}{q(\mathbf{z}_i)}dq(\mathbf{z}_i) && \text{using the iid property}\\
&= \frac{1}{N}\sum_{1}^{N}-\mathcal{L}(q,p,\mathbf{x}_i) && \text{definition of } \mathcal{L}(q,p,\mathbf{x}_i)\\
&= -\mathcal{L}(q,p,X) \triangleq -\mathcal{L}(q,p) && \text{again definition of } \mathcal{L}(p,q) \quad \square
\end{aligned}
\tag{1}
$$

normal text tiny text normal text And now to something [1] completely different ...

[1] Xifeng Guo et al. "Improved deep embedded clustering with local structure preservation.". In: *Ijcai*. 2017, pp. 1753–1759.