Possibly New Divergence

cs18m20p100002

June 20, 2022

The idea is to use a reference distribution to compare the two given distributions. With $\mathbb{P}_a$ being the reference distribution, we use $|KL(\mathbb{P}_r \| \mathbb{P}_a) - KL(\mathbb{P}_g \| \mathbb{P}_a)|$ to measure the divergence between $\mathbb{P}_r$ and $\mathbb{P}_g$. As KL-divergence is not symmetric, we also add $KL(\mathbb{P}_a \| \mathbb{P}_r) - KL(\mathbb{P}_a, \mathbb{P}_g)$ and formulate the following:

$$\begin{aligned} M(\mathbb{P}_r, \mathbb{P}_g) &= |KL(\mathbb{P}_r \| \mathbb{P}_a) - KL(\mathbb{P}_g \| P_a) + KL(\mathbb{P}_a \| \mathbb{P}_r) - KL(\mathbb{P}_a, \mathbb{P}_g)| \\ &= \left| \int_x (p_r(x) - p_g(x)) \log \frac{\sqrt{p_r(x) p_g(x)}}{p_a(x)} dx \right| \text{ where } \mathbb{P}_a = \frac{\mathbb{P}_r + \mathbb{P}_g}{2} \end{aligned} \tag{1}$$

From 1, it can be seen that $M(.,.)$ is symmetric and $M(\rho \mathbb{P}_r, \rho \mathbb{P}_g) = |\rho| \, M(\mathbb{P}_r, \mathbb{P}_g)$. Also, $M(.,.) \geq 0$ and $M(.,.) = 0 \leftrightarrow \mathbb{P}_r = \mathbb{P}_g$ following the AM-GM equality criteria. But it does not follow Triangle Inequality.

Attempt for Triangle Inequality proof to show that $M(\mathbb{P}_r, \mathbb{P}_g) \leq M(\mathbb{P}_r, \mathbb{P}_s) + M(\mathbb{P}_s, \mathbb{P}_g)$

$$\begin{aligned} M(\mathbb{P}_r, \mathbb{P}_g) &= \left| (\mathbb{P}_r - \mathbb{P}_s + \mathbb{P}_s - \mathbb{P}_g)^\top \log \frac{\sqrt{\mathbb{P}_r \mathbb{P}_g}}{\frac{\mathbb{P}_r + \mathbb{P}_g}{2}} \right| \\ &\leq \left| (\mathbb{P}_r - \mathbb{P}_s)^\top \log \frac{\sqrt{\mathbb{P}_r \mathbb{P}_g}}{\frac{\mathbb{P}_r + \mathbb{P}_g}{2}} \right| + \left| (\mathbb{P}_s - \mathbb{P}_g)^\top \log \frac{\sqrt{\mathbb{P}_r \mathbb{P}_g}}{\frac{\mathbb{P}_r + \mathbb{P}_g}{2}} \right| \\ &\leq \left| (\mathbb{P}_r - \mathbb{P}_s)^\top \log \left(\frac{\sqrt{\mathbb{P}_r \mathbb{P}_s}}{\frac{\mathbb{P}_r + \mathbb{P}_s}{2}} \sqrt{\frac{\mathbb{P}_g}{\mathbb{P}_s}} \left(\frac{\mathbb{P}_r + \mathbb{P}_s}{\mathbb{P}_r + \mathbb{P}_g} \right) \right) \right| + \left| (\mathbb{P}_s - \mathbb{P}_g)^\top \log \left(\frac{\sqrt{\mathbb{P}_g \mathbb{P}_s}}{\frac{\mathbb{P}_g + \mathbb{P}_s}{2}} \sqrt{\frac{\mathbb{P}_r}{\mathbb{P}_s}} \left(\frac{\mathbb{P}_g + \mathbb{P}_s}{\mathbb{P}_r + \mathbb{P}_g} \right) \right) \right| \end{aligned}$$

Doesn't seem to satisfy Triangle Inequality as

$$\left| (\mathbb{P}_r - \mathbb{P}_s)^\top \log \left(\sqrt{\frac{\mathbb{P}_g}{\mathbb{P}_s}} \left(\frac{\mathbb{P}_r + \mathbb{P}_s}{\mathbb{P}_r + \mathbb{P}_g} \right) \right) \right| + \left| (\mathbb{P}_s - \mathbb{P}_g)^\top \log \left(\sqrt{\frac{\mathbb{P}_r}{\mathbb{P}_s}} \left(\frac{\mathbb{P}_g + \mathbb{P}_s}{\mathbb{P}_r + \mathbb{P}_g} \right) \right) \right| > 0$$