TTIC 31230, Fundamentals of Deep Learning

David McAllester, Winter 2020

Optimal Discrimination

and Jensen-Shannon Divergence

GANs

The generator tries to fool the discriminator.

$$\Phi^* = \operatorname*{argmax}_{\Phi} \; \min_{\Psi} \; E_{\langle i, y \rangle \sim \tilde{p}_\Phi} \; -\ln P_\Psi(i|y)$$

Assuming universality of both the generator p_Φ and the discriminator P_Ψ we have $p_{\Phi^*} = \mathrm{pop}$.

Assuming Universality of Ψ Only

$$\Psi^*(\Phi) \;=\; \underset{\Psi}{\text{argmin}} \;\; E_{\langle i,\,y\rangle \sim \tilde{p}_\Phi} \;\; -\ln P_\Psi(i|y)$$

$$P_{\Psi^*(\Phi)}(1|y) \;=\; \tilde{p}_\Phi(1|y) \;=\; \frac{\tilde{p}_\Phi(1,y)}{\tilde{p}_\Phi(y)} \;=\; \frac{\frac{1}{2}\text{pop}(y)}{\frac{1}{2}\text{pop}(y) + \frac{1}{2}p_\Phi(y)}$$

$$P_{\Psi^*(\Phi)}(-1|y) \;=\; \tilde{p}_\Phi(-1|y) \;=\; \frac{\tilde{p}_\Phi(-1,y)}{\tilde{p}_\Phi(y)} \;=\; \frac{\frac{1}{2}p_\Phi(y)}{\frac{1}{2}\text{pop}(y) + \frac{1}{2}p_\Phi(y)}$$

Assuming Universality of Ψ Only

$$
\begin{aligned}
\Phi^* &= \operatorname*{argmax}_{\Phi} \; \min_{\Psi} \; E_{\langle i, y\rangle \sim \tilde{p}_\Phi} \; - \ln P_\Psi(i|y) \\
&= \operatorname*{argmax}_{\Phi} E_{(i,y) \sim \tilde{p}_\Phi} \; - \ln \tilde{p}_\Phi(i|y) \\
&= \operatorname*{argmin}_{\Phi} E_{(i,y) \sim \tilde{p}_\Phi} \; \ln \tilde{p}_\Phi(i|y) \\
&= \operatorname*{argmin}_{\Phi} \frac{1}{2} \ln \tilde{p}_\Phi(1|y) + \frac{1}{2} \ln \tilde{p}_\Phi(-1|y)
\end{aligned}
$$

Assuming Universality of Ψ Only

$$= \frac{1}{2} \ln \tilde{p}_\Phi(1|y) + \frac{1}{2} \ln \tilde{p}_\Phi(-1|y)$$

$$= \frac{1}{2} E_{y \sim \mathrm{pop}} \;\ln \frac{\frac{1}{2}\mathrm{pop}(y)}{\frac{1}{2}\mathrm{pop}(y) + \frac{1}{2}p_\Phi(y)} + \frac{1}{2} E_{y \sim p_\Phi} \;\ln \frac{\frac{1}{2}p_\Phi(y)}{\frac{1}{2}\mathrm{pop}(y) + \frac{1}{2}p_\Phi(y)}$$

$$= \frac{1}{2}\left(KL\left(\mathrm{pop}, \frac{\mathrm{pop} + p_\Phi}{2}\right) + KL\left(p_\Phi, \frac{\mathrm{pop} + p_\Phi}{2}\right)\right) - \ln 2$$

$$\Phi^* = \operatorname*{argmin}_\Phi \mathrm{JSD}(\mathrm{pop}, p_\Phi)$$

END