This handout includes space for every question that requires a written response. Please feel free to use it to handwrite your solutions (legibly, please). If you choose to typeset your solutions, the `README.md` for this assignment includes instructions to regenerate this handout with your typeset LaTeX solutions.

---

2a To help you start the proof: Using the chain rule and the fact that $\sigma'(x) = \sigma(x)(1 - \sigma(x))$,

$$\frac{\partial L_G^{\text{minimax}}}{\partial \theta} = \mathbb{E}_{\boldsymbol{z} \sim \mathcal{N}(0,I)} \left[ -\frac{\sigma'(h_\phi(G_\theta(\boldsymbol{z})))}{1 - \sigma(h_\phi(G_\theta(\boldsymbol{z})))} \frac{\partial}{\partial \theta} h_\phi\left(G_\theta\left(\boldsymbol{z}\right)\right) \right] =$$

## 3a

To help you get started with the proof: If we break the expectation up, we see that

$$\begin{aligned} L_D(\phi;\theta) &= -\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D_\phi(\boldsymbol{x})] - \mathbb{E}_{\boldsymbol{x}\sim p_\theta(\boldsymbol{x})}[\log(1-D_\phi(\boldsymbol{x}))] \\ &= -\int p_{\text{data}}(\boldsymbol{x})\log D_\phi(\boldsymbol{x})d\boldsymbol{x} - \int p_\theta(\boldsymbol{x})\log(1-D_\phi(\boldsymbol{x}))d\boldsymbol{x} \\ &= -\int \left(p_{\text{data}}(\boldsymbol{x})\log D_\phi(\boldsymbol{x}) + p_\theta(\boldsymbol{x})\log(1-D_\phi(\boldsymbol{x}))\right)d\boldsymbol{x} \\ &= \int f(D_\phi(\boldsymbol{x}))d\boldsymbol{x} \end{aligned}$$

We can set $L'_D(\phi;\theta) = 0$ to obtain the optimal $L'_D$. This yields

$$L'_D(\phi;\theta) = \frac{d}{dD_\phi(\boldsymbol{x})}\int f(D_\phi(\boldsymbol{x}))d\boldsymbol{x} = \int \frac{d}{dD_\phi(\boldsymbol{x})} f(D_\phi(\boldsymbol{x}))d\boldsymbol{x} = 0$$

Now try to apply the hint!

## 3b

To help you get started, note that

$$D_\phi(\boldsymbol{x}) = \sigma(h_\phi(\boldsymbol{x})) = \frac{1}{1+e^{-h_\phi(\boldsymbol{x})}}$$

Setting this to the expression for $D^*(\boldsymbol{x})$ in part 3a solution, we find that

## 3c

To get started

$$\begin{aligned} L_G(\theta;\phi) &= \mathbb{E}_{p_\theta(\boldsymbol{x})}[\log(1-D_\phi(\boldsymbol{x}))] - \mathbb{E}_{p_\theta(\boldsymbol{x})}[\log D_\phi(\boldsymbol{x})] \\ &= \mathbb{E}_{p_\theta(\boldsymbol{x})}\left[\log\frac{1-D_\phi(\boldsymbol{x})}{D_\phi(\boldsymbol{x})}\right] \end{aligned}$$

## 3d

## 4a To help you get started:

$$
\begin{gathered}
h_\phi(x, y) = \log \frac{p_{\mathsf{data}}(\boldsymbol{x}, y)}{p_\theta(\boldsymbol{x}, y)} \\
= \log \frac{p_{\mathsf{data}}(\boldsymbol{x}|y)}{p_\theta(\boldsymbol{x}|y)} + \log \frac{p_{\mathsf{data}}(y)}{p_\theta(y)} \\
= \log \frac{p_{\mathsf{data}}(\boldsymbol{x}|y)}{p_\theta(\boldsymbol{x}|y)} =
\end{gathered}
$$

## 5a To help you get started:

$$\mathsf{KL}(p_\theta(x) \,||\, p_{\mathsf{data}}(x)) = \mathbb{E}_{x\sim\mathcal{N}(\theta,\epsilon^2)}\left[\log\frac{\exp(-\frac{1}{2\epsilon^2}(x-\theta)^2)}{\exp(-\frac{1}{2\epsilon^2}(x-\theta_0)^2)}\right] =$$

## 5b

## 5c

## 5d