

1. 请写一下LSTM的基本公式，并介绍一下各个gates的意义。

$$\begin{aligned}\Delta c_t &= f(W_c(h_{t-1}, x_t)) \\ i_t &= f(W_i(h_{t-1}, x_t)) \\ f_t &= f(W_f(h_{t-1}, x_t)) \\ c_t &= c_{t-1}f_t + \Delta c_t i_t \\ o_t &= f(W_o(h_{t-1}, x_t)) \\ h_t &= o_t f(c_t) \\ y_t &= f(W h_t)\end{aligned}$$

- input gate i_t : 控制输入信息的进入，决定在当前时刻有哪些输入信息会被加入到状态中。
- forget gate f_t : 控制过去的信息的遗忘，决定哪些过去的信息应该被忘记。
- output gate o_t : 控制输出信息的提取，决定当前时刻哪些状态信息应该被输出。

2. 请写一下GAN模型的Loss function $V(G,D)$ ，以及其训练的min max形式。解释一下Loss的意义。

$$\begin{aligned}V(D, G) &= E_{P_{\text{data}}} [\log p_D(X)] + E_{h \sim p(h)} [\log (1 - p_D(G(h)))] \\ &= E_{P_{\text{data}}} [\log D(X)] + E_{h \sim p(h)} [\log (1 - D(G(h)))]\end{aligned}$$

训练的min max形式:

$$\min_G \max_D V(D, G)$$

该 Loss 的意义是：第一项是真实数据的对数概率，表示判别器对真实数据的判别结果；第二项是生成器生成的假数据的对数概率，表示判别器对假数据的判别结果。

3. 请证明优化GAN模型等价于优化JSD($P_{\text{data}} \parallel P_{\theta}$)。介绍一下 P_{θ} 的公式和意义——这个课堂上讲过。

$$p_{\text{mix}} = \frac{P_{\text{data}} + p_{\theta}}{2}$$

$$\begin{aligned}\text{JSD}(P_{\text{data}}|p_{\theta}) &= \text{KL}(p_{\theta}|p_{\text{mix}}) + \text{KL}(P_{\text{data}}|p_{\text{mix}}) \\ &= \sum_X \left[p_{\theta}(X) \log \frac{p_{\theta}(X)}{p_{\text{mix}}(X)} + P_{\text{data}}(X) \log \frac{P_{\text{data}}(X)}{p_{\text{mix}}(X)} \right] \\ &= -H(p_{\theta}) - H(P_{\text{data}}) - \sum_X [p_{\theta}(X) \log p_{\text{mix}}(X) + P_{\text{data}}(X) \log p_{\text{mix}}(X)] \\ &= -H(p_{\theta}) - H(P_{\text{data}}) - \sum_X \left[p_{\theta}(X) \log \frac{p_{\theta}(X)}{2(1 - D(X))} + P_{\text{data}}(X) \log \frac{P_{\text{data}}(X)}{2D(X)} \right] \\ &= -H(p_{\theta}) - H(P_{\text{data}}) + H(p_{\theta}) + H(P_{\text{data}}) + \sum_X [p_{\theta}(X) \log [2(1 - D(X))] + P_{\text{data}}(X) \log (2D(X))] \\ &= \sum_X [p_{\theta}(X) \log 2 + P_{\text{data}}(X) \log 2] + \sum_X [p_{\theta}(X) \log (1 - D(X)) + P_{\text{data}}(X) \log (D(X))] \\ &= 2 \log 2 + V(D, G)\end{aligned}$$

因此，优化GAN模型等价于优化 $\text{JSD}(P_{\text{data}}|p_{\theta})$ 。

$P_{\theta}(X)$ 的公式：

$$P_{\theta}(X) = \sum_h P(h) P_G(X|h)$$

$P_{\theta}(X)$ 表示 X 被生成出来的概率。