表示学习习题

姓名： 甘云冲

学号： 2101213081

1.

$$
\begin{aligned}
-\frac{\partial \ln P_{\boldsymbol{\theta}}(\boldsymbol{x})}{\partial \boldsymbol{\theta}} &= -\frac{\partial}{\partial \boldsymbol{\theta}} \ln \frac{\sum_{\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}}{\sum_{\boldsymbol{x},\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}} \\
&= \frac{\partial}{\partial \boldsymbol{\theta}} \ln \sum_{\boldsymbol{x},\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})} - \frac{\partial}{\partial \boldsymbol{\theta}} \ln \sum_{\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})} \\
&= \frac{1}{\sum_{\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}} \sum_{\boldsymbol{h}} \left(e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})} \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) - \frac{1}{\sum_{\boldsymbol{x},\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}} \sum_{\boldsymbol{h}} \left(e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})} \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\
&= \sum_{\boldsymbol{h}} \left(\frac{e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}}{\sum_{\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}} \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) - \sum_{\boldsymbol{h}} \left(\frac{e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}}{\sum_{\boldsymbol{x},\boldsymbol{h}} e^{-E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}} \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\
&= \sum_{\boldsymbol{h}} \left(\frac{P(\boldsymbol{x},\boldsymbol{h})}{P(\boldsymbol{x})} \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) - \sum_{\boldsymbol{h}} \left(P(\boldsymbol{x},\boldsymbol{h}) \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\
&= \sum_{\boldsymbol{h}} \left(P(\boldsymbol{h} \mid \boldsymbol{x}) \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) - \sum_{\boldsymbol{h}} \left(P(\boldsymbol{x},\boldsymbol{h}) \frac{\partial E(\boldsymbol{x},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)
\end{aligned}
$$

2.

$$
\begin{aligned}
P(h_j = 1 \mid \boldsymbol{v}) &= P(h_j = 1 \mid h_{-j}, \boldsymbol{v}) \\
&= \frac{P(h_j = 1, h_{-j}, \boldsymbol{v})}{P(h_{-j}, \boldsymbol{v})} \\
&= \frac{P(h_j = 1, h_{-j}, \boldsymbol{v})}{P(h_j = 1, h_{-j}, \boldsymbol{v}) + P(h_j = 0, h_{-j}, \boldsymbol{v})} \\
&= \frac{e^{-E(h_j=1, h_{-j}, \boldsymbol{v}; \boldsymbol{\theta})}}{e^{-E(h_j=1, h_{-j}, \boldsymbol{v}; \boldsymbol{\theta})} + e^{-E(h_j=0, h_{-j}, \boldsymbol{v}; \boldsymbol{\theta})}} \\
&= \frac{1}{1 + e^{E(h_j=1, h_{-j}, \boldsymbol{v}; \boldsymbol{\theta}) - E(h_j=0, h_{-j}, \boldsymbol{v}; \boldsymbol{\theta})}} \\
&= \frac{1}{1 + e^{-(\sum_i v_i w_{ij} + b_j)}} \\
&= \operatorname{sigmoid}(\sum_i v_i w_{ij} + b_j)
\end{aligned}
$$

$$
\begin{aligned}
P(v_i = 1 \mid \boldsymbol{h}) &= P(v_i = 1 \mid v_{-i}, \boldsymbol{h}) \\
&= \frac{P(v_i = 1, v_{-i}, \boldsymbol{h})}{P(v_{-i}, \boldsymbol{h})} \\
&= \frac{P(v_i = 1, v_{-i}, \boldsymbol{h})}{P(v_i = 1, v_{-i}, \boldsymbol{h}) + P(v_i = 0, v_{-i}, \boldsymbol{h})} \\
&= \frac{e^{-E(v_i=1, v_{-i}, \boldsymbol{h}; \boldsymbol{\theta})}}{e^{-E(v_i=1, v_{-i}, \boldsymbol{h}; \boldsymbol{\theta})} + e^{-E(v_i=0, v_{-i}, \boldsymbol{h}; \boldsymbol{\theta})}} \\
&= \frac{1}{1 + e^{E(v_i=1, v_{-i}, \boldsymbol{h}; \boldsymbol{\theta}) - E(v_i=0, v_{-i}, \boldsymbol{h}; \boldsymbol{\theta})}} \\
&= \frac{1}{1 + e^{-(\sum_j w_{ij} h_j + a_i)}} \\
&= \mathrm{sigmoid}(\sum_j w_{ij} h_j + a_i)
\end{aligned}
$$