Exercise Sheet 6

Machine Learning 2, SS16

June 2, 2016

Mario Tambos, 380599; Viktor Jeney, 348969; Sascha Huk, 321249; Jan Tinapp, 0380549

Exercise 1

(a)
Define

$$y := x^T W - b^T$$

$$\begin{aligned}
p_\theta(x) &= \sum_{h \in \{-1,0,1\}^N} p(x,h) \\
&= \sum_{h \in \{-1,0,1\}^N} \frac{1}{Z} \exp(yh + x^T a) \\
&= \frac{1}{Z} \exp(x^T a) \sum_{h \in \{-1,0,1\}^N} \exp(yh) \\
&= \frac{1}{Z} \exp(x^T a) \sum_{h \in \{-1,0,1\}^N} \exp(\sum_{i=1}^{N} y_i h_i) \\
&= \frac{1}{Z} \exp(x^T a) \sum_{h \in \{-1,0,1\}^N} \prod_{i=1}^{N} \exp(y_i h_i)
\end{aligned}$$

Because the expression $\exp(y_i h_i)$ only depends on the i'th component of h, we can rewrite the sum and product to get:

$$\begin{aligned}
p_\theta(x) &= \frac{1}{Z} \exp(x^T a) \prod_{i=1}^{N} \sum_{h \in \{-1,0,1\}} \exp(y_i h_i) \\
&= \frac{1}{Z} \exp(x^T a) \exp(\log(\prod_{i=1}^{N} \sum_{h \in \{-1,0,1\}} \exp(y_i h_i))) \\
&= \frac{1}{Z} \exp(x^T a) \exp(\sum_{i=1}^{N} \log(\sum_{h \in \{-1,0,1\}} \exp(y_i h_i))) \\
&= \frac{1}{Z} \exp(x^T a) \exp(\sum_{i=1}^{N} \log(1 + e^{y_i} + e^{-y_i})) \\
&= \frac{1}{Z} \exp(x^T a) \exp(\sum_{i=1}^{N} \log(1 + 2cosh(y_i))) \\
&= \frac{1}{Z} \exp(x^T a + \sum_{i=1}^{N} \log(1 + 2cosh(w_i x - b_i)))
\end{aligned}$$

$$= \frac{1}{Z} \exp(x^T a + \sum_{i=1}^{N} \log(2(\frac{1}{2} + cosh(w_i x - b_i))))$$

$$= \frac{1}{Z} \exp(x^T a + \sum_{i=1}^{N} \log(2) + (\frac{1}{2} + cosh(w_i x - b_i)))$$

$$= \frac{1}{Z} \exp(N \log(2) + x^T a + \sum_{i=1}^{N} (\frac{1}{2} + cosh(w_i x - b_i)))$$

$$= \frac{1}{Z} 2^N \exp(x^T a + \sum_{i=1}^{N} (\frac{1}{2} + cosh(w_i x - b_i)))$$

With

$$Z' := \frac{2^N}{Z}$$

the desired result follows.

(b)

First compute gradients of F:

$$\bigtriangledown_a F$$