Exercise Sheet 6

Machine Learning 2, SS16

June 2, 2016

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

Exercise 1 - RBM with Ternary Hidden Units

Model: $(x_i \perp\!\!\!\perp x_j | h)$ and $(h_i \perp\!\!\!\perp h_j | x)$ for all $i \neq j$.

(a) We have to show:

$$\sum_{h \in \{-1,0,1\}^N} e^{x^\top a + x^\top W h + h^\top b} \quad \propto \quad e^{x^\top a + \sum_{j=1}^N \log(\frac{1}{2} + \cosh(w_j^\top x + b_j))}$$

Proof. Devide both sides by $\exp(x^\top a)$ and write down the multiplication by h more explicitly:

$$\sum_{h \in \{-1,0,1\}^N} e^{\sum_{j=1}^N h_j w_j^\top x + h_j b_j} \quad \propto \quad e^{\sum_{j=1}^N \log(\frac{1}{2} + \cosh(w_j^\top x + b_j))}$$

It follows by laws of exponentiation:

$$\sum_{h \in \{-1,0,1\}^N} \prod_{j=1}^N e^{h_j w_j^\top x + h_j b_j} \quad \propto \quad \cancel{\exp(\log} \prod_{j=1}^N (\frac{1}{2} + \cosh(w_j^\top x + b_j))$$

Because the expression $\exp(h_j w_j^\top x + h_j b_j)$ only depends on the j'th component of h, we can rewrite the sum and product to get:

$$\prod_{j=1}^N \sum_{h_j \in \{-1,0,1\}} e^{h_j w_j^\top x + h_j b_j} \quad \propto \quad \prod_{j=1}^N \frac{(1 + 2\cosh(w_j^\top x + b_j))}{2}$$

Unfold the three cases on the left side and pull out the constant factor on the right side:

$$\prod_{j=1}^N \left(e^{w_j^\top x + b_j} + e^{-w_j^\top x - b_j} + 1 \right) \quad \propto \quad \frac{1}{2^N} \prod_{j=1}^N (1 + 2\cosh(w_j^\top x + b_j))$$

Use definition of cosh on the left side:

$$\prod_{j=1}^N \left(2\cosh(w_j^\top x + b_j) + 1 \right) \quad \propto \quad \frac{1}{2^N} \prod_{j=1}^N (1 + 2\cosh(w_j^\top x + b_j))$$

□

(c)
UNDER CONSTRUCTION!
$\Pr(x) = \frac{1}{Z}\prod_{j=1}^{N} \exp(-W_j x + a^\top x - b_j) + \exp(W_j x + a^\top x + b_j) + \exp(a^\top x)$
$\Pr(x) = \frac{1}{Z}\exp(a^\top x)\prod_{j=1}^{N} \exp(-W_j x - b_j) + \exp(W_j x + b_j) + 1$

$$\Pr(x) = \sum_{h\in\{-1,0,1\}^N} e^{x^\top a + x^\top W h + h^\top b}$$

$$\Pr(h_j = -1 \mid x) = \frac{\Pr(h_j = -1, x)}{\Pr(x)} = \frac{\frac{1}{Z}\prod_{j=1}^{N} \exp(-W_j x + a^\top x - b_j)}{\Pr(x)} = \prod_{j=1}^{N} \frac{1}{1 + \exp(2W_j x + 2b_j) + \exp(W_j x + b_j)}$$

$$\Pr(h_j = 1 \mid x) = \frac{\Pr(h_j = 1, x)}{\Pr(x)} = \frac{\frac{1}{Z}\prod_{j=1}^{N} \exp(+W_j x + a^\top x + b_j)}{\Pr(x)} = \prod_{j=1}^{N} \frac{1}{1 + \exp(-2W_j x - 2b_j) + \exp(-W_j x - b_j)}$$

$$\Pr(h_j = 0 \mid x) = \frac{\Pr(h_j = 0, x)}{\Pr(x)} = \frac{\frac{1}{Z}\prod_{j=1}^{N} \exp(a^\top x)}{\Pr(x)} = \prod_{j=1}^{N} \frac{1}{1 + \exp(-W_j x - b_j) + \exp(W_j x + b_j)}$$

$$\Pr(x_i = 0 \mid h) = \frac{\Pr(x_i = 0, h)}{\Pr(h)} =$$