

Exercise Sheet 12

1 Deep SVDD

a) Minimum at $\vec{w} = \vec{0}$, $b = -1$

$$\begin{aligned} J &= \min_w \frac{1}{N} \sum_{i=1}^N \| \vec{w}^T x_i - 1 \|^2 = \min_w \frac{1}{N} \sum_{i=1}^N (\vec{w}^T x_i - 1)(\vec{w}^T x_i - 1) \\ &= \min_w \frac{1}{N} \sum_{i=1}^N \vec{w}^T x_i \cdot x_i^T \vec{w} - 2 \vec{w}^T x_i + 1 \\ &= \min_w \vec{w}^T \frac{\sum_{i=1}^N x_i x_i^T}{N} \vec{w} - \frac{2}{N} \sum_{i=1}^N x_i^T \vec{w} + \frac{1}{N} \sum_{i=1}^N 1 \\ \frac{\partial J}{\partial \vec{w}} &= (\Sigma + \Sigma^T) \vec{w} - 2 \frac{1}{N} \sum_{i=1}^N x_i = 0 \end{aligned}$$

$$\Leftrightarrow 2 \Sigma \vec{w} - 2 \bar{x} = 0 \quad \Leftrightarrow \vec{w} = \Sigma^{-1} \bar{x}$$

2 Restricted Boltzmann Machine

$$\begin{aligned} a) \quad p(h_k = 1 | x) &= \frac{p(x, h_k = 1)}{p(x)} = \frac{\sum_{h_{-k}} p(x, h_k = 1, h_{-k})}{p(x)} \\ &= \frac{\sum_{h_{-k}} p(x, h_k = 1, h_{-k})}{\sum_{q \in \{0,1\}} \sum_{h_{-k}} p(x, h_k = q, h_{-k})} \\ &= \frac{\sum_{h_{-k}} \frac{1}{2} \exp(x^T w_{:,k} + h_k - E(x, h_{-k}))}{\sum_{q \in \{0,1\}} \sum_{h_{-k}} \frac{1}{2} \exp(x^T w_{:,k} q + h_k q - E(x, h_{-k}))} \\ &= \frac{\exp(x^T w_{:,k} + h_k) \sum_{h_{-k}} \exp(-E(x, h_{-k}))}{\sum_{q \in \{0,1\}} \exp(x^T w_{:,k} q + h_k q) \sum_{h_{-k}} \exp(-E(x, h_{-k}))} \\ &= \frac{\exp(x^T w_{:,k} + h_k)}{1 + \exp(x^T w_{:,k} + h_k)} = \text{sigmoid}(x^T w_{:,k} + h_k) \end{aligned}$$

b)

$$\begin{aligned}
 p(x_j = 1|h) &= \frac{p(x_j = 1, h)}{p(h)} = \frac{\sum_{x_i} p(x_j = 1, x_i, h)}{\sum_{q \in \{0,1\}} \sum_{x_i} p(x_j = q, x_i, h)} \\
 &= \frac{\sum_{x_i} \frac{1}{2} \exp(w_{ji}^T h - E(x_i, h))}{\sum_{q \in \{0,1\}} \sum_{x_i} \frac{1}{2} \exp(w_{ji}^T h - E(x_i, h))} \\
 &= \frac{\exp(w_{ji}^T h)}{1 + \exp(w_{ji}^T h)} = \text{sigm}(w_{ji}^T h)
 \end{aligned}$$

c)

$$\begin{aligned}
 p(x) &= \sum_h p(x, h) = \sum_h \frac{1}{2} \exp(-E(x, h)) \\
 &= \sum_h \frac{1}{2} \exp(x^T W h - b^T h) = \sum_h \frac{1}{2} \exp\left(\sum_k x^T W_{:,k} h_k - b_k^T h_k\right) \\
 &= \frac{1}{2} \sum_k \pi_k \exp(x^T W_{:,k} h_k - b_k^T h_k) \\
 &= \frac{1}{2} \pi_k (1 + \exp(x^T W_{:,k} - b_k)) \\
 &= \frac{1}{2} \exp(\log(\pi_k (1 + \exp(x^T W_{:,k} - b_k))))
 \end{aligned}$$