# Task 1

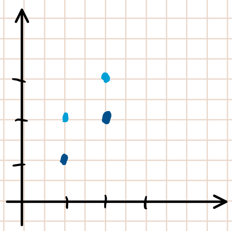
a) $\sigma = \frac{1}{1+e^{-x}}$

$$\frac{d\sigma}{dx} = -e^{-x} \cdot (-1) \frac{1}{1+e^{-x}} = \frac{e^{-x}}{(1+e^{-x})^2}$$

b) $$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1} = \frac{1-e^{-2x}}{1+e^{-2x}} = \frac{2-(1+e^{-2x})}{1+e^{-2x}} = 2\left(\frac{1}{1+e^{-2x}}\right)-1$$

$$= 2\sigma(2x) - 1$$

c)



$x_1^1 = (1,1)$ , $x_2^1 = (2,2)$

$x_1^2 = (1,2)$ , $x_2^2 = (2,3)$

$w^T x + b > 0$ for class 1

$w^T x + b < 0$ for class 2

$$\Rightarrow \begin{array}{ll} \mathrm{I} & w_1 + w_2 + b > 0 \\ \mathrm{II} & 2w_1 + 2w_2 + b > 0 \\ \mathrm{III} & w_1 + 2w_2 + b < 0 \\ \mathrm{IV} & 2w_1 + 3w_2 + b < 0 \end{array}$$

$\mathrm{III} - \mathrm{I} \Rightarrow w_2 < 0$ choose $w_2 = -1$

$$\Rightarrow \begin{array}{ll} \mathrm{I} & w_1 - 1 + b > 0 \\ \mathrm{II} & 2w_1 - 2 + b > 0 \\ \mathrm{III} & w_1 - 2 + b < 0 \\ \mathrm{IV} & 2w_1 - 3 + b < 0 \end{array} \Longleftrightarrow \begin{array}{ll} \mathrm{I} & w_1 + b > 1 \\ \mathrm{II} & 2w_1 + b > 2 \\ \mathrm{III} & w_1 + b < 2 \\ \mathrm{IV} & 2w_1 + b < 2 \end{array}$$

$\mathrm{I}, \mathrm{III} \Rightarrow 1 < w_1 + b < 2$

$\mathrm{II}, \mathrm{IV} \Rightarrow 2 < 2w_1 + b < 3$

choose $w_1 = 1$, $b = .5$

$$\Rightarrow w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad b = .5$$

## Task 3

a) $z^1$ and $z^2$ are identical, bc they have a const. offset

(i) $$\mathrm{softmax}(z + c\mathbb{1}; \lambda)_k = \frac{\exp[\lambda(z_k + c)]}{\sum_j \exp[\lambda(z_j + c)]} = \frac{\exp(\lambda z_k)\exp(\lambda c)}{\exp(\lambda c)\sum_j \exp(\lambda z_j)}$$
$$= \mathrm{softmax}(z; \lambda)_k$$

(ii) $$\mathrm{softmax}(cz; \lambda)_k = \frac{\exp(\lambda c z_k)}{\sum_j \exp(\lambda c z_j)} = \mathrm{softmax}(z; c\lambda)_k \neq \mathrm{softmax}(z; \lambda)_k$$

d) $$\frac{\partial \mathrm{lse}(z; \lambda)}{\partial z_k} = \frac{\partial}{\partial z_k}\left[\frac{1}{\lambda}\log\left(\sum_j \exp(\lambda z_j)\right)\right] = \frac{1}{\cancel{\lambda}} \frac{1}{\sum_j \exp(\lambda z_j)} \cancel{\lambda}\exp(\lambda z_k)$$
$$= \mathrm{softmax}$$

e) $$\lim_{\lambda \to \infty} = \mathrm{lse}(z; \lambda) = \lim_{\lambda \to \infty} \frac{1}{\lambda}\log\left[\sum_j \exp(\lambda z_j)\right], \quad z_m = \max_j z_j$$
$$= \lim_{\lambda \to \infty} \frac{1}{\lambda}\log\left[\exp(\lambda z_m)\sum_j \exp(\lambda(z_j - z_m))\right]$$
$$= \lim_{\lambda \to \infty} \frac{1}{\lambda}\left[\lambda z_m + \log \sum_j \exp(\lambda(z_j - z_m))\right]$$
$$= z_m + \lim_{\lambda \to \infty} \frac{1}{\lambda} \log \underbrace{\sum_j \exp(\lambda(z_j - z_m))}_{\leq k}$$
$$= z_m = \max(z)$$