a)

$$l(w) = \sum_{n=0}^{N-1} ln(1 + e^{-y_n w^T x_n})$$

Where $w \in R^p, y \in \{-1; 1\}, x \in R^p$

Find gradient w.r.t. "w" : $\frac{dl(w)}{dw}$

My try is:

$$\begin{bmatrix} \frac{dl(w)}{dw_1} \\ \frac{dl(w)}{dw_2} \\ \dots \\ \frac{dl(w)}{dw_n} \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^{N-1} \frac{-y_n x_n e^{-y_n w^T x_n}}{1+e^{-y_n w^T x_n}} x_1 \\ \sum_{n=0}^{N-1} \frac{-y_n x_n e^{-y_n w^T x_n}}{1+e^{-y_n w^T x_n}} x_2 \\ \dots \\ \sum_{n=0}^{N-1} \frac{-y_n x_n e^{-y_n w^T x_n}}{1+e^{-y_n w^T x_n}} x_i \end{bmatrix}$$

b)

$$l(w) = \sum_{n=0}^{N-1} ln(1 + e^{-y_n w^T x_n}) + C w^T w$$

where

$$C \geq 0$$

the regularization strength parameter

Task is the same: find gradient of regularized log-loss function