Homework 4
78572 - Telma Correia
78958 - Gonçalo Rodrigues

(a)

$$
\begin{aligned}
w^* &= argmin_w\{-log \prod_{n=1}^{\infty} P(a_n|x_n, w)\} \\
&= argmin_w\{-\sum_{n=1}^{\infty} log(P(a_n|x_n, w))\} \\
&= argmin_w\{-\sum_{n=1}^{\infty} log(\pi(a_n|x_n, w))\} \\
&= argmin_w\{-\sum_{n=1}^{\infty} a_n log(\pi(1|x_n, w) + (1-a_n) log(\pi(0|x_n, w)\} \\
&= argmin_w\{-\sum_{n=1}^{\infty} a_n log(\pi(1|x_n, w) + (1-a_n) log(1-\pi(1|x_n, w)\} \\
&= argmax_w\{\sum_{n=1}^{\infty} a_n log(\pi(1|x_n, w) + (1-a_n) log(1-\pi(1|x_n, w)\}
\end{aligned}
$$

(b) Denoting $\sigma(x) = \frac{1}{1+e^{-x}}$ as the sigmoid function, and its derivative as $\sigma'(x) = \sigma(x)(1-\sigma(x))$, then we have that:

$$\frac{d}{dw}\pi(1|x_n, w) = \frac{d}{dw}\sigma(w^T x_n) = (\frac{d}{dw} w x_n)\sigma(w^T x_n)(1-\sigma(w^T x_n)) = x_n \pi(1|x_n, w)(1-\pi(1|x_n, w))$$

We'll be using the above result in exercise (b) and (c)

$$
\begin{aligned}
\frac{d}{dw}l(D,\pi) &= (\sum_{n=1}^{N} a_n log(\pi(1|x_n,w) + (1-a_n)log(1-\pi(1|x_n,w))' \\
&= \sum_{n=1}^{N}(a_n log(\pi(1|x_n,w))' + ((1-a_n)log(1-\pi(1|x_n,w))' \\
&= \sum_{n=1}^{N} a_n log'(\pi(1|x_n,w))(\pi(1|x_n,w))' + (1-a_n)log'(1-\pi(1|x_n,w))(1-\pi(1|x_n,w))' \\
&= \sum_{n=1}^{N} a_n \frac{1}{\pi(1|x_n,w)} x_n \pi(1|x_n,w)(1-\pi(1|x_n,w)) - \\
&(1-a_n)\frac{1}{1-\pi(1|x_n,w)} x_n \pi(1|x_n,w)(1-\pi(1|x_n,w)) \\
&= \sum_{n=1}^{N} a_n x_n(1-\pi(1|x_n,w)) - (1-a_n)x_n\pi(1|x_n,w) \\
&= \sum_{n=1}^{N} a_n x_n - a_n x_n \pi(1|x_n,w) - x_n\pi(1|x_n,w) + a_n x_n \pi(1|x_n,w) \\
&= \sum_{n=1}^{N} a_n x_n - x_n\pi(1|x_n,w) \\
&= \sum_{n=1}^{N} x_n(a_n - \pi(1|x_n,w))
\end{aligned}
$$

(c)

$$
\begin{aligned}
H &= \frac{d}{dw} g \\
&= (\sum_{n=1}^{N} x_n (a_n - \pi(1|x_n, w))' \\
&= \sum_{n=1}^{N} x_n ((a_n - \pi(1|x_n, w))' \\
&= \sum_{n=1}^{N} -x_n (\pi(1|x_n, w))' \\
&= -\sum_{n=1}^{N} x_n (\pi(1|x_n, w))' \\
&= -\sum_{n=1}^{N} x_n x_n^T (\pi(1|x_n, w))(1 - \pi(1|x_n, w))
\end{aligned}
$$