1.1 Sol: In the trianing set,when computing $(X^T X)^{-1}$, the matrix should not have non-zero rows in echleon form.i.e., if $N \geq D + 1$ and training set is linearly independent of each other we get unique solution.

1.2 Sol:

RSS(ω) $= \sum_n [y_n - (b + \sum_d \omega_d x_{nd})]^2$

derivative w.r.t to b and equating to 0

$$\partial RSS(\omega)/\partial b = 2 \sum_n [y_n \text{ - (b + } \sum_d \omega_d x_{nd})](\text{-1}) = 0$$

$$\sum_n y_n - Nb - \sum_n \sum_d \omega_d x_{nd} = 0$$

$$Nb = \sum_n y_n - \sum_n \sum_d \omega_d x_{nd}$$

but, $\frac{1}{N} \sum_n x_{nd} = 0$

$$b^* = \frac{\sum_n y_n}{N}$$

1.3 Sol:

As we do not have access to feature x , implies x_n =0, implies ω=0

$$\partial \varepsilon(b)/\partial b = -\sum_n \{y_n[1 - \sigma(b)] - (1 - y_n)\sigma(b)\} = 0$$

$$\sigma(b) = \frac{1}{N} \sum_n y_n => e^{-b} = \frac{N}{\sum_n y_n} - 1, assuming \sum_n 1 = N$$

$$\boxed{b^* = \log(\frac{\sum_n y_n}{N - \sum_n y_n})}$$

probability that y=1 is $p(y = 1) = \frac{\sum_n y_n}{N}$