Training the model

$$p(y_{1:N} \mid x_{1:N}, \underline{w}) \; ; \quad \underline{w} \sim p(\underline{w})$$

$$p(\underline{w} \mid x_{1:N}, y_{1:N}) \propto p(y_{1:N} \mid x_{1:N}, \underline{w}) \, p(\underline{w})$$

NOT ANALYTICALLY AV.

Max. A Posteriori Estimate of $\underline{w}$:

$$\max_{\underline{w}} \, \log p(y_{1:N} \mid x_{1:N}, \underline{w}) + \log p(\underline{w})$$

$$= \sum_{i=1}^{N} \log p(y_i \mid x_i, \underline{w}) + \underbrace{\log p(\underline{w})}_{0}$$

$$= \sum_{i=1}^{N} \Big\{ y_i \log \operatorname{sigm}(w_0 + w_1 x_i) + (1 - y_i)\big[1 - \operatorname{sigm}(w_0 + w_1 x_i)\big] \Big\}$$