$$L(w) = \prod p_w(x_i, y_i)$$

$$= \prod p_w(y_i \mid x_i)\, p(x_i)$$

$$y_i = f_w(x_i) + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma)$$

$$y_i \mid x_i \overset{D}{=} N(f_w(x_i), \sigma).$$

$$\propto \prod_{i=1}^{N} \left[\exp\left(- \frac{(y_i - f_w(x_i))^2}{2} \right) p(x_i) \right]$$

$$Y_i = f_\omega(x_i) + \varepsilon_i$$

$$Y_i = \mathbb{1}\{ f_\omega(x_i) + \varepsilon_i \geq 0 \}.$$

$$L(\omega) = \prod_{i} p_\omega(x_i, y_i).$$

$$= \prod_{i=1} p_\omega(y_i \mid x_i)\, p(x_i)$$

$$Y_i \mid X_i = \mathrm{Ber}(q(x_i))$$

$$p_\omega(y_i \mid x_i) = \mathbb{P}(Y_i = y_i \mid X_i = x_i)$$

$$q(x_i) = \mathbb{P}(Y_i = 1 \mid X_i = x_i)$$

$$= \mathbb{P}(f_\omega(x_i) + \varepsilon_i \geq 0)$$

$$= \int_{\mathbb{R}} \mathbb{1}\{ f_\omega(x_i) + g \geq 0 \} \times \varphi(g)\, dg.$$

where φ is the density of $N(0,1)$.

$$L(w) = \prod^{N} p_w(x_i, y_i).$$

$$= \prod_{i=1} p_w(y_i \mid x_i)\, p(x_i)$$

$$q(x_i) = \mathbb{P}(Y_i = 1 \mid X_i = x_i)$$

$$z_i \overset{iid}{\sim} \mathcal{B}(\theta) \qquad \theta \in (0, 1).$$

$$L^z(\theta \mid (z_i)) = \prod \theta^{z_i} (1 - \theta)^{1 - z_i}$$

Applying the analogy with

$\theta \leftrightarrow q(x_i)$ and $z_i \leftarrow y_i$.

we get

$$L(w) = \prod_{i=1}^{N} \left\{ q(x_i)^{y_i} (1 - q(x_i))^{1 - y_i} \times p(x_i) \right]$$

$$\prod \theta^{z_i} (1 - \theta)^{1 - z_i} = \theta^{\sum z_i} (1 - \theta)^{N - \sum z_i}$$

$$L(w) = \prod_{i} p_w(x_i, y_i).$$

$$= \prod_{i=1} p_w(y_i | x_i) \, p(x_i)$$

$$q(x_i) = \mathbb{P}(Y_i = 1 \mid X_i = x_i)$$

$$p_w(y_i | x_i) = q(x_i)^{y_i} \, (1 - q(x_i))^{1-y_i}$$

Bayes formula.

- $p(x, y)$ a complete likelihood
- $p(x|y) \, p(y)$

$$p(y|x) = \frac{p(x|y) \, p(y)}{p(x)}.$$

Bernoulli model

$Y_i \sim \mathrm{Ber}(q) \qquad q \in (0,1)$

$$L(q \mid (y_i)) = \prod q^{y_i} (1-q)^{1-y_i} .$$

Beta distribution on $(0,1)$
with density

$$t \longmapsto \frac{t^{\alpha-1} (1-t)^{\beta-1}}{B(\alpha,\beta)} \qquad \alpha,\beta > 0 .$$
$\in (0,1)$

Then posterior

$$p(q \mid (y_i)) \propto L(q \mid (y_i))\, p_{prior}(q)$$

$$\propto \left[\prod q^{y_i} (1-q)^{1-y_i}\right] q^{\alpha-1} (1-q)^{\beta-1}$$

$$\propto q^{\sum y_i + \alpha - 1} (1-q)^{N - \sum y_i + \beta - 1} .$$

$$= \mathrm{Beta}\left(\sum y_i + \alpha,\ N - \sum y_i + \beta\right) .$$

$$y_i = f_\omega(x_i) + \sigma^2 \varepsilon_i \qquad \varepsilon_i \overset{iid}{\sim} N(0,1)$$

$$(y_i)_{i=1}^N = N_n\left(\begin{pmatrix} f_\omega(x_1) \\ \vdots \\ f_\omega(x_n) \end{pmatrix}, \beta^{-1} I_n\right)$$

$$\beta^{-1} = \sigma^2.$$

$$Y_i = f_\omega(X_i) + \sigma \varepsilon_i \qquad \varepsilon_i \overset{iid}{\sim} N(0,1).$$

$$\omega \mapsto p_{prior}(\omega)$$

$$p(\omega \mid D) = ?$$

$$L(\omega; D) = \prod \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(y_i - f_\omega(x_i))^2}{2\sigma^2}\right).$$

$$p(\omega \mid D) \propto L(\omega; D)\, p_{prior}(\omega).$$

in the case $p_{prior}(\omega) = N(m_0, S_0)$

$$p(\omega \mid D) \propto \exp\Big(-\|y - \Phi_x \omega\|^2 \frac{\beta}{2}$$

$$- \frac{1}{2} \langle S_0^{-1}(\omega - m_0), \omega - m_0 \rangle\Big)$$

We want to show that

$p(\omega \mid D)$ is Gaussian.

We show

$$p(w \mid D) \propto \exp\left(-w^T A w + b^T w\right)$$

$$\propto \exp\left(-\frac{1}{2} \langle S_N^{-1}(w - m_N), w - m_N \rangle\right)$$

where

$$S_N^{-1} = 2A$$

$$S_N^{-1} m_N = b$$

$$p(w \mid D) \propto \exp\left(-\|y - \Phi_x w\|^2 \frac{\beta}{2} - \frac{1}{2} \langle S_0^{-1}(w - m_0), w - m_0 \rangle\right)$$

$$\propto \exp\left(-\frac{1}{2} \langle (\beta \Phi_x^T \Phi_x + S_0^{-1}) w, w \rangle + \langle \beta \Phi_x^T y + S_0^{-1} m_0, w \rangle\right).$$

$$S_N^{-1} = \beta \Phi_x^T \Phi_x + S_0^{-1}$$

$$m_N = S_N \left(\beta \Phi_x^T y + S_0^{-1} m_0 \right).$$