

Advanced probabilistic methods - Sketch

Christian Segercrantz 481056

April 5, 2022

$$-\log p(\mathcal{D} \mid \mathbf{W}) = -\log \prod_{i=0}^1 \mathcal{N}(y_i \mid w_0 + w_1 x_i, \sigma_l^2) \quad (1)$$

$$= \sum_{i=0}^N -\log \mathcal{N}(y_i \mid w_0 + w_1 x_i, \sigma_l^2) \quad (2)$$

$$-\log \mathcal{N}(y_i \mid w_0 + w_1 x_i, \sigma_l^2) = -\log \frac{1}{\sigma_l \sqrt{2\pi}} + \frac{1}{2\sigma_l^2} (y_i - w_0 - w_1 x_i)^2 \quad (3)$$

$$= \log \sigma_l \sqrt{2\pi} + \frac{1}{2\sigma_l^2} (y_i - w_0 - w_1 x_i)^2 \quad (4)$$

$$-\log p(\mathcal{D} \mid \mathbf{W}) = \sum_{i=0}^N \log \sigma_l \sqrt{2\pi} + \frac{1}{2\sigma_l^2} \sum_{i=0}^N (y_i - w_0 - w_1 x_i)^2 \quad (5)$$

$$= N \log \sigma_l \sqrt{2\pi} + \frac{1}{2\sigma_l^2} \sum_{i=0}^N (y_i - w_0 - w_1 x_i)^2, \quad \log \sigma_l \sqrt{2\pi} \approx 0.001 \quad (6)$$

$$\approx \frac{1}{2\sigma_l^2} \sum_{i=0}^N (y_i - w_0 - w_1 x_i)^2 \quad (7)$$

Since the prior $p(\mathbf{w})$ is a MVN with diagonal covariance and that the mean-field approximation of the posterior $q(\mathbf{w})$ is product of gaussians we can write the KL-divergence as a sum of the separate

parts

$$\text{KL}[q(\mathbf{w})|p(\mathbf{w})] = \text{KL}_0[q(w_0)|p(w_0)] + \text{KL}_1[q(w_1)|p(w_1)] \quad (8)$$

$$\text{KL}_i[q(w_i) | p(w_i)] = \int -q(w_i) \log \frac{p(w_i)}{q(w_i)} dw_i \quad (9)$$

$$= \underbrace{\int q(w_i) \log q(w_i) dw_i}_{\text{entropy} = -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2}} - \int q(w_i) \log p(w_i) dw_i \quad (10)$$

$$q(w_i) = \mathcal{N}(w_i | \mu_i, \sigma_i^2) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_i^2}(w_i - \mu_i)^2\right), \quad \sigma_i \geq 0 \quad (11)$$

$$p(w_i) = \mathcal{N}(w_i | 0, \alpha) = \frac{1}{\alpha \sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2} w_i^2\right), \quad \alpha \geq 0 \quad (12)$$

$$\log p(w_i) = -\log(\alpha \sqrt{2\pi}) - \frac{1}{2\alpha^2} w_i^2 \quad (13)$$

$$\text{KL}_i[q(w_i) | p(w_i)] = -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} - \int q(w_i) \log p(w_i) dw_i \quad (14)$$

$$\text{KL}_i[q(w_i) | p(w_i)] = -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} - \int \mathcal{N}(w_i | \mu_i, \sigma_i^2) \left(-\log(\alpha \sqrt{2\pi}) - \frac{1}{2\alpha^2} w_i^2\right) dw_i \quad (15)$$

$$= -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} + \log(\alpha \sqrt{2\pi}) \underbrace{\int \mathcal{N}(w_i | \mu_i, \sigma_i^2) dw_i}_{=1} + \frac{1}{2\alpha^2} \int \mathcal{N}(w_i | \mu_i, \sigma_i^2) w_i^2 dw_i \quad (16)$$

$$= -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} + \log(\alpha \sqrt{2\pi}) + \frac{1}{2\alpha^2} \underbrace{\int \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_i^2}(w_i - \mu_i)^2\right) w_i^2 dw_i}_{\mu_i^2 + \sigma_i^2} \quad (17)$$

$$= -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} + \log(\alpha \sqrt{2\pi}) + \frac{\mu_i^2 + \sigma_i^2}{2\alpha^2} \quad (18)$$

$$= -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} + \frac{1}{2} \log(2\pi\alpha^2) + \frac{\mu_i^2 + \sigma_i^2}{2\alpha^2} \quad (19)$$

$$= -\frac{1}{2} \log(2\pi\sigma_i^2) - \frac{1}{2} + \frac{1}{2} \log(2\pi\alpha^2) + \frac{\mu_i^2 + \sigma_i^2}{2\alpha^2} \quad (20)$$

$$= -\frac{1}{2} + \log\left(\frac{\alpha}{\sigma_i}\right) + \frac{\mu_i^2 + \sigma_i^2}{2\alpha^2} \quad (21)$$

$$(22)$$