

# Week 4 - Regularized Linear Regression

NGUYEN NHU QUYNH

October 2021

---

## Bayes Theorem

- We have:

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

$$\Leftrightarrow \text{posterior} = \frac{\text{likelihood.prior}}{\text{evidence}}$$

$$\Rightarrow P(w|x, t, \alpha, \beta) = \frac{P(t|x, w, \beta).P(w|\alpha)}{P(x, t, \alpha, \beta)}$$

- Otherwise:

$$P(w|\alpha) \sim N(0, \alpha^{-1}I)$$

$$\Rightarrow P(t|x, w, \beta) \sim N(t|y(x, w), \beta^{-1})$$

To maximize posterior, we maximize  $P(t|x, w, \beta).P(w, \alpha)$  and using maximum likelihood.

$$\Rightarrow P(t|x, w, \beta).P(w, \alpha) = \prod_{i=1}^N N(t_i|y(x_i, w_i), \beta^{-1})N(0, \alpha^{-1}I)$$

$$= \frac{1}{\sqrt{2\pi\beta^{-1}}} \exp\left(\frac{-\beta}{2}(t_n - y(x_n - w))^2\right) \frac{1}{(2\pi)^D |\alpha^{-1}I|} \exp\left(\frac{-1}{2}w^T(\alpha^{-1}I)^{-1}w\right)$$

$$\log P(t|x, w, \beta).P(w|\alpha) = \frac{-\beta}{2} \sum_{i=1}^N (y(x_i, w) - t_i)^2 - \frac{\alpha}{2} w^T w$$

- Alternatively, we can minimize the negative log-likelihood.

$$NLL(t|x, w, \beta) = \frac{1}{2}\beta \sum_{i=1}^N (y(x, w) - t)^2 + \frac{\alpha}{2} w^T w$$

$$\Rightarrow S_{min} = \|Xw - t\|^2 + \lambda w^T w$$

with  $\lambda = \frac{\alpha}{\beta}$

$$S_{min} = (Xw - t)^T (Xw - t) + \lambda w^T w$$

$$S_{min} = t^T t - 2(Xw)^T t + (Xw)^T (Xw) + \lambda w^T w$$

$$S_{min} = t^T t - 2w^T X^T t + w^T X^T X w + \lambda w^T w$$

$$\text{With } \frac{dS}{dw^T} = -2X^T t + 2X^T X w + 2\lambda I w = 0$$

$$\Rightarrow 2(X^T X + \lambda I)w = 2X^T t$$

$$\Leftrightarrow w = (X^T X + \lambda I)^{-1} X^T t$$